

FP6 CEDER Project Deliverable 2.1

“Design of systems to be used for pilot”

Software architecture of the CEDER project's
pilot systems

N. Dayagi (Corr.), E. Rosenberg (Corr.),
M. Grønfeldt (Sirius), T. Hjörvar (FRI), A. Kideys (JRC), U. Kröner (JRC),
J. Neves (NEAFC), F. Quirijns (IMARES), G. Pétursdóttir (FRI),
G. Pilling (Cefas), D. Reid (FRS), K. Sünksen (GINR)



EUR 24019 EN - 2009

The mission of the JRC-IPSC is to provide research results and to support EU policy-makers in their effort towards global security and towards protection of European citizens from accidents, deliberate attacks, fraud and illegal actions against EU policies.

European Commission
Joint Research Centre
Institute for the Protection and Security of the Citizen

Contact information

Address: TP 051, Joint Research Centre, Via E. Fermi 2749, 21027 Ispra (VA), Italy
E-mail: Ulrich.Kroener@jrc.ec.europa.eu
Tel.: +39 0332 78 6719
Fax: +39 0332 78 9658

<http://ipsc.jrc.ec.europa.eu/>
<http://www.jrc.ec.europa.eu/>

Legal Notice

Neither the European Commission nor any person acting on behalf of the Commission is responsible for the use which might be made of this publication.

***Europe Direct is a service to help you find answers
to your questions about the European Union***

**Freephone number (*):
00 800 6 7 8 9 10 11**

(*) Certain mobile telephone operators do not allow access to 00 800 numbers or these calls may be billed.

A great deal of additional information on the European Union is available on the Internet.
It can be accessed through the Europa server <http://europa.eu/>

JRC 54607

EUR 24019 EN
ISBN 978-92-79-13739-6
ISSN 1018-5593
DOI 10.2788/40321

Luxembourg: Office for Official Publications of the European Communities

© European Communities, 2009

Reproduction is authorised provided the source is acknowledged

Printed in Italy

Contents

Contents	3
1 Introduction	5
1.1 Executive Summary	5
1.2 Context.....	5
2 JRC CEDER TAC and Quota Prototype System	6
2.1 Top Level Design.....	6
2.2 General.....	6
2.2.1 The System Design	6
2.2.2 Web Site.....	8
2.2.3 Web Services.....	19
2.3 The database	36
2.3.1 Introduction	36
2.3.2 PostGres JRC Reference Tables	36
2.3.3 PostGres JRC Input Tables	37
2.3.4 PostGres JRC Output Tables.....	38
3 Icelandic Prototype System	40
3.1 Introduction	40
3.1.1 Purpose	40
3.1.2 Scope.....	40
3.2 System Overview.....	41
3.3 Data Collection Module	41
3.3.1 Import VMS-positions, catch and landing data.....	41
3.3.2 VMS-positions CSV-file format.....	42
3.3.3 Catch and landings CVS-file format.....	42
3.4 Analysis Module	43
3.4.1 Track Classification and Effort estimation	43
3.4.2 Catch Prediction.....	43
3.5 Presentation Module.....	43
3.5.1 View vessel tracks.....	43
3.5.2 View fleet activity	43
3.5.3 View predicted catch.....	43
3.6 Database design.....	44
3.6.1 Entity-relationship diagram	44
4 CEDER System for Greenland fisheries	45
4.1 Introduction	45
4.2 Database Design.....	45
4.2.1 General.....	45
4.3 System design.....	45
4.3.1 Input sources of the GFLK database	45
4.3.2 Design of database for catch-control and Quota-management.....	46
4.3.3 Catch and quota calculation.....	47
4.3.4 Data-mining	47
5 Correlation Systems: CEDER Web Site.....	48
5.1 Top Level Design.....	48
5.1.1 General.....	48
5.1.2 The System Design	48
5.2 The Web Site – Detailed Design	50
5.2.1 Input Analyses	50
5.2.2 Parameters Update	52
5.2.3 Processing Analyses	55
5.2.4 Query Analyses	57
5.2.5 Statistics	58

(Page intentionally left blank)

1 Introduction

1.1 Executive Summary

Addressing the uncertainties in fishing activities, the FP6 CEDER project designed a set of pilot systems to demonstrate the feasibility of providing information in a timelier manner.

This document describes the software designs of the aforementioned pilot systems. The developers of these systems made choices mainly regarding front-ends, databases, import/export of files, and web services. As such this document is a collection of technical “Software Architecture Documents” describing CEDER’s pilot systems.

1.2 Context

The objective of WP2.1 is to design pilot systems that can demonstrate the feasibility of bringing real-time information on fisheries to stakeholders.

Although reporting obligations are similar in most of the nations participating in this project, there are national variations in the systems used. Therefore it is not possible to develop a one-size-fits-all project. Therefore separate prototypes were developed:

1. for Icelandic fisheries led by DIS
2. for Greenland fisheries led by Sirius IT
3. for UK fisheries led by Correlation Systems

JRC developed a TAC and quota prediction prototype, which was implemented as a separated prototype system.

Deliverable 2.1 is a report outlining the design of the system to be used for the trials, for each prototype system that was developed. Note that the basic algorithms have been developed in WP1.2, therefore, here the report includes for each prototype system decisions on a software framework to encapsulate these algorithms. It includes development of input procedures compatible with data streams of VMS, landing declarations and electronic logbooks coming from the fisheries of this project.

2 JRC CEDER TAC and Quota Prototype System

By: Ulrich Kröner, JRC

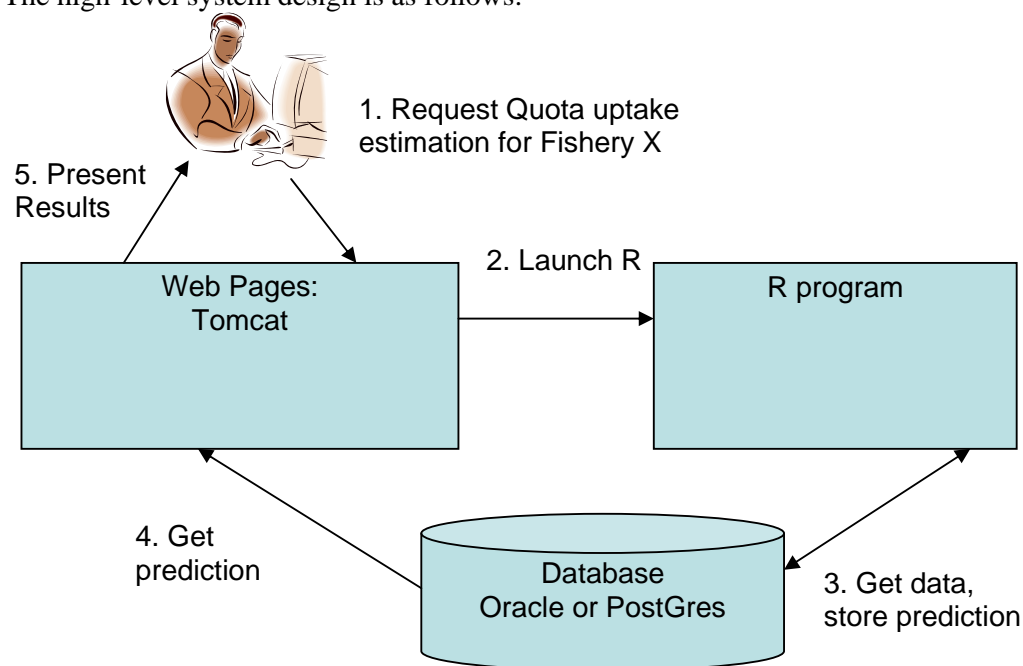
2.1 Top Level Design

The CEDER TAC and quota prediction prototype system is implemented as a web site. The web site enables the user to find the best landings and/or quota uptake time series model. A future extension to the system will include web services.

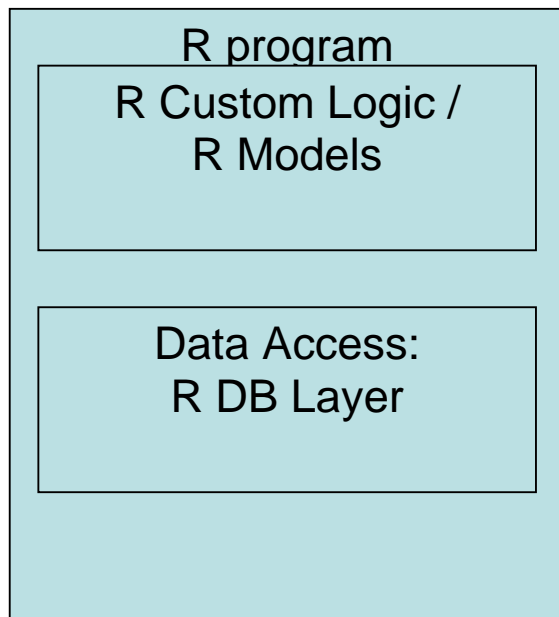
2.2 General

2.2.1 The System Design

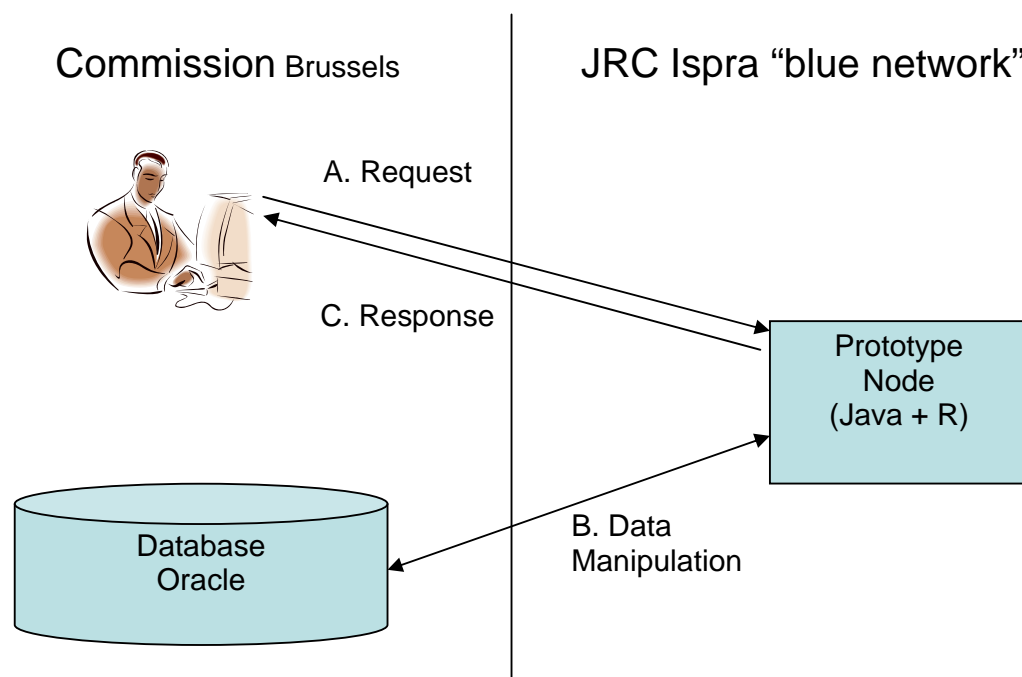
The high-level system design is as follows:



Concerning the R program, we used a layered approach, separating database access logic from the program's algorithms:



The system would be deployed as follows:



We chose the following node for the main installation of the prototype system:
D01RI0700130.net1.cec.eu.int

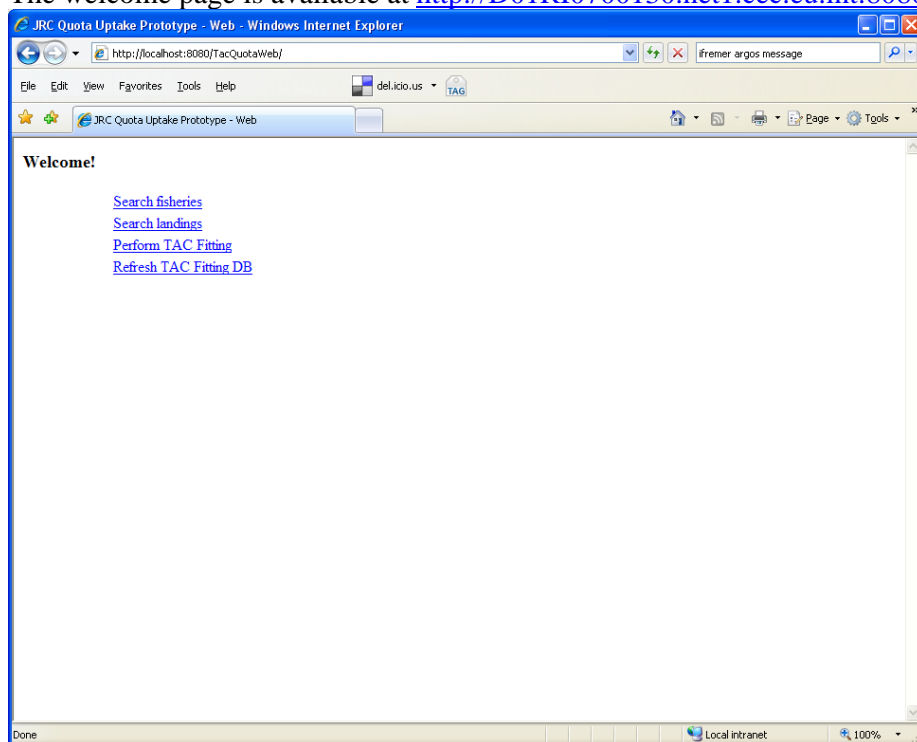
2.2.2 Web Site

In the documentation on the web site, it is assumed that the system is installed on D01RI0700130.net1.cec.eu.int and that one has access via the European Commission Intranet.

Nevertheless, the prototype can also be installed on other systems. In that case, when using the system according to the documentation laid down below, the host name of “D01RI0700130.net1.cec.eu.int” needs to be replaced by the appropriate host name.

2.2.2.1 Welcome Page

The welcome page is available at <http://D01RI0700130.net1.cec.eu.int:8080/TacQuotaWeb>



From there, one can query for fisheries defined in the database, or query for landings copied from FIDES CRONT, or one can launch the TAC and Quota Prediction R process, or one can refresh the prototype's database .

On the Fisheries and Landings screens, we use the “query by example” rules. This means that

- If you leave a field blank, it selects all matching database rows.
For example, you can leave “species” blank to extend your query to all species in the database.
- If you enter a value in a field, then it selects those rows that match the field’s value. Matching is approximate and case-insensitive.
For example, you can enter 2006 into the year field, and this will limit the search to rows pertaining to 2006. If you enter the cipher “2” into an area field, then this will match areas “2AC4.” and “F27”, but not “07.”.
- If you complete various fields, then it searches for rows that fulfil both criteria, rather than searching for rows that fulfil either criterion¹.
For example, if you both enter ANF for the species and 2006 for the year, it searches for rows from 2006, which also concern Anglerfish.

¹ It uses SQL “AND” semantics; it is not possible to perform an “OR”.

2.2.2.2 Fishery Page

On <http://D01RI0700130.net1.cec.eu.int:8080/TacQuotaWeb/pages/fisheryPage.jsp> one finds the fishery page. For instance, one can find all fisheries defined that are both British and have a “07”-ish area code. Then one presses the button at the bottom, and gets:

View Fisheries - Windows Internet Explorer

http://localhost:8080/TacQuotaWeb/fisheryAction.do

File Edit View Favorites Tools Help

View Fisheries

View Fisheries

Fill out the form, then press submit. You will then be able to see all fisheries that fit the entries of the form.

CATCH-COUNTRY CODE

SPECIES

AREA CODE

FISHERY

REGIME CODE

Country	Species	Area	Regime	Fishery	Comment
GBR	ANF	07.	COCR-QS	NSAF	UK Anglerfish in all of ICES zone 7, 2a comm waters, 4 comm waters, 5 6 12 and 14
GBR	COD	07.	COCR-QS	NSR	UK Cod in FAO 27, ICES 4, all of ICES 4, 2a, ...
GBR	HAD	07.	COCR-QS	NSR	UK Haddock in FAO 27, ICES 4, all of ICES 4, 2a, ...
GBR	PLE	07.	COCR-QS	NSF	UK and NL Plaice in all of zone 7
GBR	WHG	07.	COCR-QS	NSR	UK and NL Whiting in FAO 27, ICES 4, all of ICES 4, 2a, ...
GBR	COD	07A.	COCR-QS	NSR	UK Cod in FAO 27, ICES 4, all of ICES 4, 2a, ...
GBR	HAD	07A.	COCR-QS	NSR	UK Haddock in FAO 27, ICES 4, all of ICES 4, 2a, ...
GBR	PLE	07A.	COCR-QS	NSF	UK and NL Plaice in all of zone 7
GBR	SOL	07A.	COCR-QS	NSF	UK Sole in all of zone 7
GBR	WHG	07A.	COCR-QS	NSR	UK and NL Whiting in FAO 27, ICES 4, all of ICES 4, 2a, ...
GBR	SOL	07D.	COCR-QS	NSF	UK Sole in all of zone 7
GBR	SOL	07E.	COCR-QS	NSF	UK Sole in all of zone 7

Done Local intranet 100%

2.2.2.3 Landing Page

On <http://D01RI0700130.net1.cec.eu.int:8080/TacQuotaWeb/pages/landingPage.jsp> one finds the landing page. For instance, one can find all landings defined that are French, are for the year 2005, and apply to big eye tuna:

View Landings

Fill out the form, then press submit. You will then be able to see all landings that fit the entries of the form.

YEAR: 2005
 SPECIES CODE: BET
 AREA CODE:
 CATCH-COUNTRY CODE: FRA

Country	Species	Area	Year	Month	Cumulative Kgs	Adjusted Quota (uptake pct)	Current Quota (uptake pct)
FRA	BET	ATLANT	2005	1	122.0	9438.0 (1%)	9438.0 (1%)
FRA	BET	ATLANT	2005	2	202.0	9438.0 (2%)	9438.0 (2%)
FRA	BET	ATLANT	2005	3	313.0	9438.0 (3%)	9438.0 (3%)
FRA	BET	ATLANT	2005	4	365.9	9438.0 (4%)	9438.0 (4%)
FRA	BET	ATLANT	2005	5	453.6	9438.0 (5%)	9438.0 (5%)
FRA	BET	ATLANT	2005	6	622.0	9438.0 (7%)	9438.0 (7%)
FRA	BET	ATLANT	2005	7	624.8	9438.0 (7%)	9438.0 (7%)
FRA	BET	ATLANT	2005	8	992.3	9438.0 (11%)	9438.0 (11%)
FRA	BET	ATLANT	2005	9	1217.5	9438.0 (13%)	9438.0 (13%)
FRA	BET	ATLANT	2005	10	1231.9	9438.0 (13%)	9438.0 (13%)
FRA	BET	ATLANT	2005	11	1328.5	9438.0 (14%)	9438.0 (14%)
FRA	BET	ATLANT	2005	12	1400.0	9438.0 (15%)	9438.0 (15%)

2.2.2.4 TAC Prediction Query Page

The query page for TAC and quota prediction is available at <http://D01RI0700130.net1.cec.eu.int:8080/TacQuotaWeb/pages/tacFittingPage.jsp>.

When pointing one's browser to the above URL, the server returns a form:

TAC and Quota Time Series Fitting

Fill out the form, then press submit. Statistical analysis may take a while.

Mandatory

Species: ANF

Area: 07

Catch Country Code: GBR

Optional

Start from year (0: include all years): 0

Up to year and month (format: YYYYMM): 999999

Ignore data up to month (threshold from which "quota uptake" makes sense): 2

Predict how many months: 3

Confidence level: 0.6

How many models perform predictions: 5

This page acts fundamentally different from the fishery and landings pages, in that exact values need to be entered for the 3 mandatory arguments.

Tip: In order to get possible tuples for country, area, and species, the user is advised to enter values as returned by the fishery page.

Mandatory arguments are:

Name	Type	Description
country	ISO values	The country for the batch run, select an ISO country code (ex: 'GBR')
area	FIDES area values	The area for the batch run, select an area known to FIDES, which is also in the database (ex: '04-N.').
species	ISO values	The species for the batch run, select an ISO species (ex: 'ANF')

Optional arguments are:

Name	Type	Description
------	------	-------------

Start From	Year on 4 ciphers	Indicates that data should only be considered at and after this year (ex: '2002' to consider only data from January 2002 included.). The default is to start from the first year in the database that is found for the given species, country, and area (equivalent to '0')
Finish at year and month	Year on 4 ciphers, followed by month on 2 ciphers	Indicates that data should only be considered up to the given year and month. (Ex: '200610' to consider only data to October 2006 included.) The default is to consider all data found in the database for the given species, country, and area (equivalent to '999999').
Minimal month required	Positive integer, a month, usually < 5	Minimal months required at the beginning of a year. (Ex: '2'). At the beginning of every year, quotas are evaluated. In predicting fisheries quota uptake, a prediction that tries to predict quota uptake at the beginning of a year, makes little sense. The default is to disregard data up to and including February, corresponding to '2'.
Predict number of months	Positive integer, number of months, usually < 5	How many months should the prediction attempt to predict (ex: '3'). By default, this is worth '3', so the program tries to tune the models for 3 months of lookahead, and then tries to predict the next 3 months from the current data.
Confidence level	Positive fraction of 1	Indicates the prediction confidence level to be used when predicting values. This only works on some models, for instance linear models. By default, it will be set to 0.6, which corresponds to “more likely than not”.
Number of winning models	Positive integer, size of “podium”	The program determines winning models according to a fitness criterion. Winning models will then perform predictions, which the user can see. This parameter controls how many winners there are. (Ex: '5').

By pushing the button, one launches the statistical process. While the process is running, one sees the following screen.

Time Series Fitting - Windows Internet Explorer

http://localhost:8080/TacQuotaWeb/tacFittingAction.do

uniform distribution

File Edit View Favorites Tools Help

Time Series Fi... W Standard error (...) W Normal distributi... W Confidence inter... ABC of Data Pro...

TAC and Quota Time Series Fitting

Fill out the form, then press submit. Statistical analysis may take a while.

Mandatory

Species

Area

Catch Country Code

Optional

Start from year (0: include all years)

Up to year and month (format: YYYYMM)

Ignore data up to month
(threshold from which "quota uptake" makes sense)

Predict how many months

Confidence level

How many models perform predictions

TAC fitting process is still running. This form will reload in 10 seconds.

Process output

Process error output

The screen refreshes every once in a while, displaying the ongoing output of the R process. The total duration varies, but should not exceed a few minutes.

The large text fields will fill up with the output of the R session. These are useful in case of problems, for diagnostics purposes.

2.2.2.5 TAC Prediction Results Page

Once the prediction process is finished, the user can see the results page, composed of graphs and time series model information:

TAC and Quota Time Series Fitting

Fill out the form, then press submit. Statistical analysis may take a while.

Mandatory	
Species	ANF
Area	07.
Catch Country Code	GBR
Optional	
Start from year (0: include all years)	0
Up to year and month (format: YYYYMM)	999999
Ignore data up to month (threshold from which "quota uptake" makes sense)	2
Predict how many months	3
Confidence level	0.6
How many models perform predictions	5

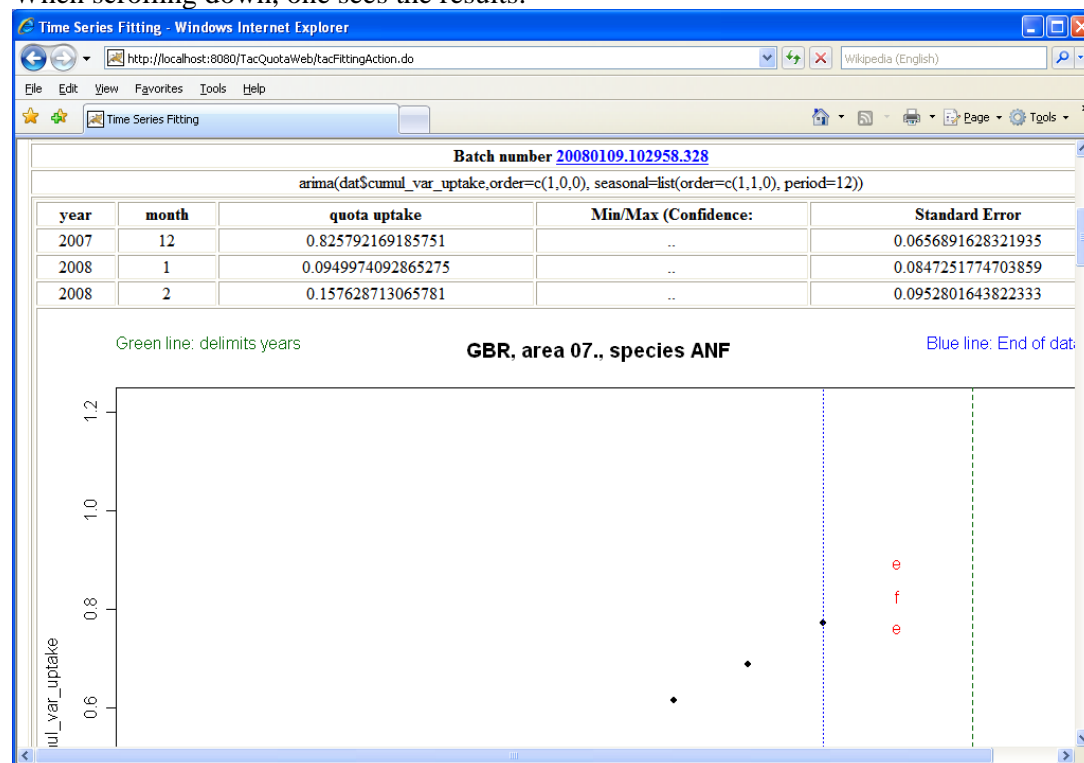
R Formula

Batch number [20080109.102958.328](#)

`arima(dat$cumul_var_uptake,order=c(1,0,0), seasonal=list(order=c(1,1,0), period=12))`

Note: If one wants to see the same results again without having to re-run the process, one can bookmark the “Batch number” link.

When scrolling down, one sees the results:



The graph represents a visual of the data (black) and the prediction (red). The letter “f” stands for the fitted prediction, the “e” stands for the predicted standard error.

The contents of the table are to be interpreted as follows:

- The model which won the “model fitting contest” was one where the cumulative variable quota uptake² is modelled as a seasonally adjusted ARIMA. This ARIMA first differentiates the data from one year to the next³, then regresses the difference between the same months of different years⁴ to an AR(1) process, and then regresses any difference left from one month to the next of the same year⁵ to another AR(1) process.
- The winning model guesses that the quota uptake in December 2007 is 82.6% of the total allocated quota at that time.
- For December 2006, the model predicts a standard error of 6.57%, or a standard deviation of about 25%.

Note: This is a very large standard deviation, which is basically not yielding any information. But the model was “elected” because of its quota uptake guesses, not because of the standard error predictions.

Typically, one of the winning models is a linear model, usually one needs to scroll down a bit. These models provide confidence intervals.

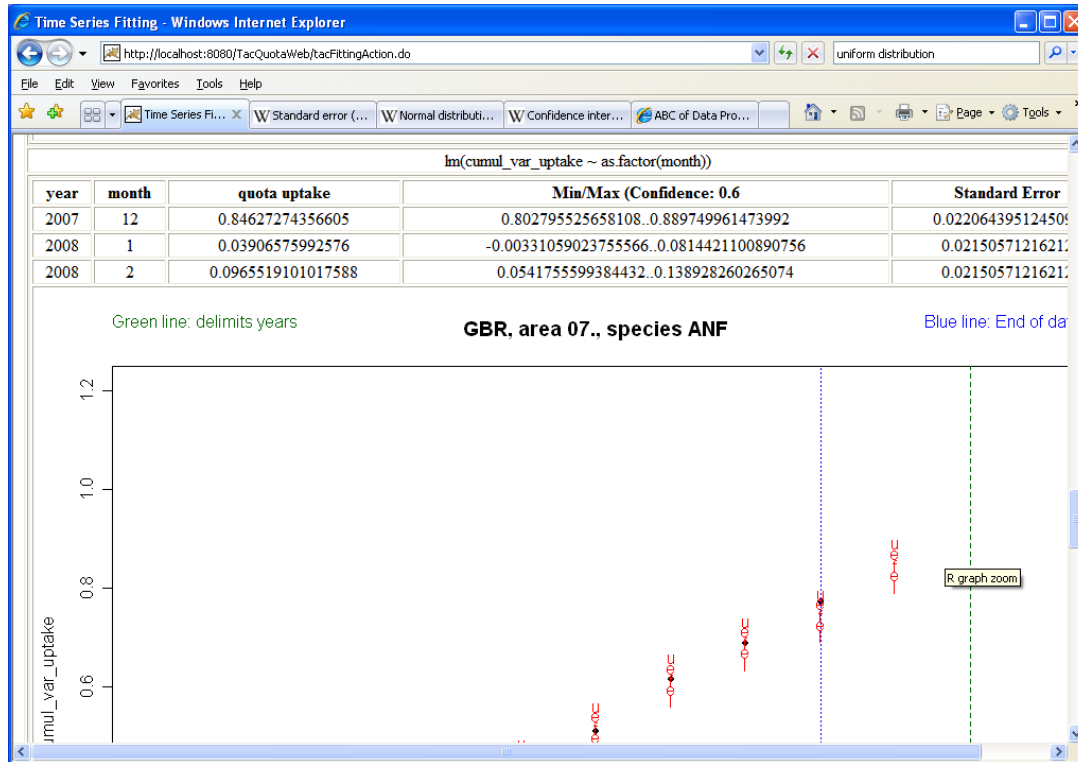
An example is depicted below:

² cumul_var_uptake: Defined as part of quota consumed at a given month, as part of the total quota, which was valid at that moment in time.

³ order = (1, 0, 0), seasonal = list(order(1, 1, 0))

⁴ order = (1, 0, 0), seasonal = list(order(1, 1, 0))

⁵ order = (1, 0, 0), seasonal = list(order(1, 1, 0))



The red letters “u” and “l” are the upper and lower bounds of the 60% confidence interval.

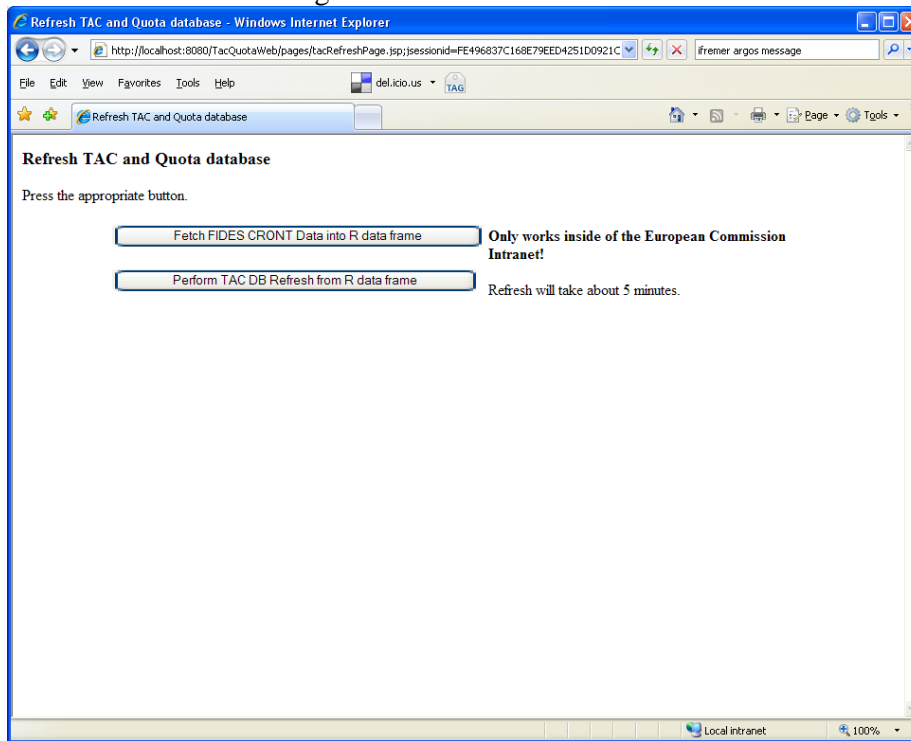
The contents of the table are to be interpreted as follows:

- A linear model, where the cumulative variable uptake depends on the month of year (“variable quota uptake looks the same every year”), seems to be doing reasonably well (because it is in the winning models from the contest)
- The model guesses that the quota uptake in December 2007 is 84.6% of the total allocated quota at that time.
- With 60% confidence, the model predicts that the December 2007 value is between 80% and 90%.

At the bottom of the page, you find the output of the R process. This exists mainly for the case in which the R process terminates abnormally, so that the user can send feedback to the author.

2.2.2.6 Refresh page

By clicking the last link on the welcome page, one can see the page used to refresh the prototype's database. It contains 2 large buttons.



The FIDES CRONT database contains landings per area, country, and target species. It is accessible only from within the European Commission Intranet.

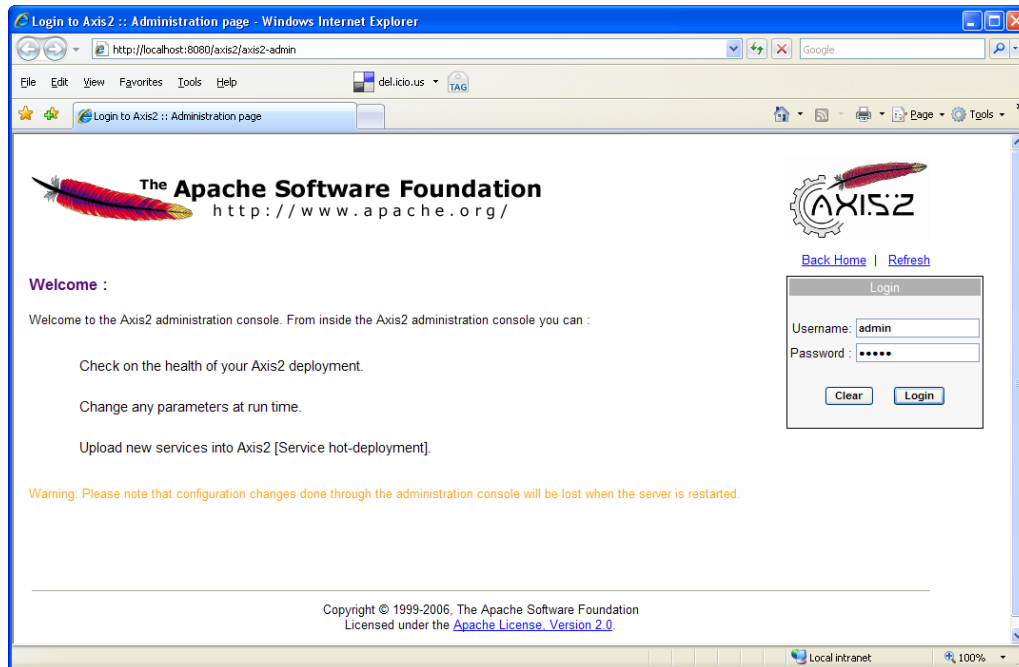
Please run the two tasks in sequence. Push the upper button first, then wait until the browser displays the results. If the first task was successful, you will be informed via a textual message on the result page ("Info: Success!"). Then push the lower button, and again wait until the browser displays the results.

But why do we first fetch the data into some intermediary format, and then refresh the prototype's DB from the intermediary format? This is done because it enables the system to be used "off site", meaning on a computer not located at the European Commission.

In case you plan to run an instance of the prototype system "off site", you will in addition a distribution mechanism for the R data files. We have such a mechanism in place, so please contact us on receiving the proper R data files.

2.2.3 Web Services

We developed Web Services for all of the functionalities available on the web pages. We used the Axis2 software package, version 1.3. This software package has an administration page, accessible using the URL <http://D01RI0700130.net1.cec.eu.int:8080/axis2/axis2-admin>. The password is axis2.



After logging in, one reaches the Axis2 main page. More information on Axis2 can be found at <http://ws.apache.org/axis2/1.3/contents.html>

We developed one service group, called TacQuotaWeb. It contains a set of services, of which 5 shall be documented:

1. FisheryService: Queries the fishery table.
2. LandingService: Queries the landing table, updates the effort figures.
3. TacCmdLineService: Performs TAC and Quota Prediction.
4. TacFidesCrontLoadService: Reads relevant contents from the FIDES CRONT database, dumps them into R data files. Requires that one uses prototype instance running on the Commission's network.
5. TacRefreshInputTablesService: Using the contents of the R data files refreshes the prototype's tables.

All of these services should only be invoked using the SOAP v1.1 binding.

Some of the web pages make use of the “query by example” rules. The web services are no different.

The Fishery and Landing services use the “query by example” rules. This means that

- If you leave a field blank, it selects all matching database rows.
For example, you can leave “species” blank to extend your query to all species in the database.
- If you enter a value in a field, then it selects those rows that match the field's value. Matching is approximate and case-insensitive.
For example, you can enter 2006 into the year field, and this will limit the search to rows

pertaining to 2006. If you enter the cipher “2” into an area field, then this will match areas “2AC4.” and “F27”, but not “07.”.

- If you complete various fields, then it searches for rows that fulfil both criteria, rather than searching for rows that fulfil either criterion⁶.
For example, if you both enter ANF for the species and 2006 for the year, it searches for rows from 2006, which also concern Anglerfish.

⁶ It uses SQL “AND” semantics; it is not possible to perform an “OR”.

2.2.3.1 Fishery Service

The fishery service is located at <http://D01RI0700130.net1.cec.eu.int:8080/axis2/services/FisheryService> and is described by a WSDL, which can be obtained by adding the string “?wsdl” at the end of the above URL.

A SOAP client can transmit the following SOAP request:

```
<soapenv:Envelope
  xmlns:soapenv="http://schemas.xmlsoap.org/soap/envelope/"
  xmlns:q0="http://ws.ceder.jrc.it/xsd"
  xmlns:q1="http://gen.bean.tac.fishreg.jrc.it/xsd"
  xmlns:xsd="http://www.w3.org/2001/XMLSchema"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance">
  <soapenv:Body>
    <q0:getFisheryArray>
      <q0:sample>
        <q1:area>2AC4.</q1:area>
        <q1:country>NLD</q1:country>
      </q0:sample>
    </q0:getFisheryArray>
  </soapenv:Body>
</soapenv:Envelope>
```

The above example is requesting all fisheries for area 2AC4. and country NLD.

The server replies with the following SOAP response:

```
<soapenv:Envelope
  xmlns:soapenv="http://schemas.xmlsoap.org/soap/envelope/">
  <soapenv:Body>
    <getFisheryArrayResponse
      xmlns:ax29="http://gen.bean.tac.fishreg.jrc.it/xsd"
```

```
xmlns:ns="http://ws.ceder.jrc.it/xsd">
<ns:return type="it.jrc.fishreg.tac.bean.gen.Fishery">
  <ax29:area>2AC4.</ax29:area>
  <ax29:comment>
    UK and NL Plaice in all of zone 7
  </ax29:comment>
  <ax29:country>NLD</ax29:country>
  <ax29:fishery>NSF</ax29:fishery>
  <ax29:iid>1129</ax29:iid>
  <ax29:regime>COCR-QS</ax29:regime>
  <ax29:species>PLE</ax29:species>
</ns:return>
<ns:return type="it.jrc.fishreg.tac.bean.gen.Fishery">
  <ax29:area>2AC4.</ax29:area>
  <ax29:comment>
    NL Cod in FAO 27, ICES 4, all of ICES 4, 2a, ...
  </ax29:comment>
  <ax29:country>NLD</ax29:country>
  <ax29:fishery>NSR</ax29:fishery>
  <ax29:iid>1152</ax29:iid>
  <ax29:regime>COCR-QS</ax29:regime>
  <ax29:species>COD</ax29:species>
</ns:return>
<ns:return type="it.jrc.fishreg.tac.bean.gen.Fishery">
  <ax29:area>2AC4.</ax29:area>
  <ax29:comment>
    NL Haddock in FAO 27, ICES 4, all of ICES 4, 2a, ...
  </ax29:comment>
  <ax29:country>NLD</ax29:country>
  <ax29:fishery>NSR</ax29:fishery>
  <ax29:iid>1163</ax29:iid>
  <ax29:regime>COCR-QS</ax29:regime>
  <ax29:species>HAD</ax29:species>
</ns:return>
<ns:return type="it.jrc.fishreg.tac.bean.gen.Fishery">
  <ax29:area>2AC4.</ax29:area>
  <ax29:comment>
    UK and NL Whiting in FAO 27, ICES 4, all of ICES 4,
```

```

                2a, ...
            </ax29:comment>
            <ax29:country>NLD</ax29:country>
            <ax29:fishery>NSR</ax29:fishery>
            <ax29:iid>1173</ax29:iid>
            <ax29:regime>COCR-QS</ax29:regime>
            <ax29:species>WHG</ax29:species>
        </ns:return>
    </getFisheryArrayResponse>
</soapenv:Body>
</soapenv:Envelope>

```

This means that the species PLE, COD, HAD, and WHG are present.

2.2.3.2 Landing Service

The landing service is located at <http://D01RI0700130.net1.cec.eu.int:8080/axis2/services/LandingService> and is described by a WSDL, which can be obtained by adding the string “?wsdl” at the end of the above URL.

Get Landings

A SOAP client can transmit the following SOAP request to get the landings:

```

<soapenv:Envelope
  xmlns:soapenv="http://schemas.xmlsoap.org/soap/envelope/"
  xmlns:q0="http://ws.ceder.jrc.it/xsd"
  xmlns:q1="http://gen.bean.tac.fishreg.jrc.it/xsd"
  xmlns:xsd="http://www.w3.org/2001/XMLSchema"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance">
  <soapenv:Body>
    <q0:getLandingArray>
      <q0:sample>
        <q1:area>2AC4.</q1:area>
        <q1:country>NLD</q1:country>
        <q1:species>PLE</q1:species>
        <q1:year>2004</q1:year>
      </q0:sample>
    </q0:getLandingArray>
  </soapenv:Body>
</soapenv:Envelope>

```

```

        </q0:sample>
      </q0:getLandingArray>
    </soapenv:Body>
  </soapenv:Envelope>

```

The above example is requesting all landings for species PLE, country NLD, year 2004 and area “2AC4.”.

The server replies with the following SOAP response, which we have abbreviated:

```

<soapenv:Envelope
  xmlns:soapenv="http://schemas.xmlsoap.org/soap/envelope/">
  <soapenv:Body>
    <getLandingArrayResponse
      xmlns:ax28="http://gen.bean.tac.fishreg.jrc.it/xsd"
      xmlns:ns="http://ws.ceder.jrc.it/xsd">
      <ns:return type="it.jrc.fishreg.tac.bean.gen.Landing">
        <ax28:area>2AC4.</ax28:area>
        <ax28:country>NLD</ax28:country>
        <ax28:cumulKgs>3173.6</ax28:cumulKgs>
        <ax28:effort>0.0</ax28:effort>
        <ax28:kgs>3173.6</ax28:kgs>
        <ax28:landingId>15039</ax28:landingId>
        <ax28:month>1</ax28:month>
        <ax28:quotaFinal>23599.0</ax28:quotaFinal>
        <ax28:quotaVar>22649.0</ax28:quotaVar>
        <ax28:species>PLE</ax28:species>
        <ax28:year>2004</ax28:year>
      </ns:return>
    </getLandingArrayResponse>
  </soapenv:Body>
</soapenv:Envelope>

```

... etc ...

```

<ns:return type="it.jrc.fishreg.tac.bean.gen.Landing">
  <ax28:area>2AC4.</ax28:area>
  <ax28:country>NLD</ax28:country>
  <ax28:cumulKgs>5373.9</ax28:cumulKgs>
  <ax28:effort>0.0</ax28:effort>
  <ax28:kgs>2200.3</ax28:kgs>
  <ax28:landingId>15040</ax28:landingId>
  <ax28:month>2</ax28:month>
  <ax28:quotaFinal>23599.0</ax28:quotaFinal>

```



```

        <ax28:quotaVar>22649.0</ax28:quotaVar>
        <ax28:species>PLE</ax28:species>
        <ax28:year>2004</ax28:year>
    </ns:return>
    <ns:return type="it.jrc.fishreg.tac.bean.gen.Landing">
        <ax28:area>2AC4.</ax28:area>
        <ax28:country>NLD</ax28:country>
        <ax28:cumulKgs>23665.7</ax28:cumulKgs>
        <ax28:effort>0.0</ax28:effort>
        <ax28:kgs>2191.8</ax28:kgs>
        <ax28:landingId>15050</ax28:landingId>
        <ax28:month>12</ax28:month>
        <ax28:quotaFinal>23599.0</ax28:quotaFinal>
        <ax28:quotaVar>23599.0</ax28:quotaVar>
        <ax28:species>PLE</ax28:species>
        <ax28:year>2004</ax28:year>
    </ns:return>
</getLandingArrayResponse>
</soapenv:Body>
</soapenv:Envelope>

```

These are the landing figures for the requested year, for months January to December

Update effort figure

A SOAP client can transmit the following SOAP request to update the associated effort figures:

```

<soapenv:Envelope
  xmlns:soapenv="http://schemas.xmlsoap.org/soap/envelope/"
  xmlns:q0="http://ws.ceder.jrc.it/xsd"
  xmlns:q1="http://gen.bean.tac.fishreg.jrc.it/xsd"
  xmlns:xsd="http://www.w3.org/2001/XMLSchema"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance">
  <soapenv:Body>
    <q0:updateEffort>
      <q0:sample>
        <q1:area>07.</q1:area>
        <q1:country>GBR</q1:country>

```

```

        <q1:effort>42</q1:effort>
        <q1:month>11</q1:month>
        <q1:species>ANF</q1:species>
        <q1:year>2004</q1:year>
      </q0:sample>
    </q0:updateEffort>
  </soapenv:Body>
</soapenv:Envelope>

```

This updates the effort value to 42, for the entry matching country GBR, species ANF, area “07.”, year 2004, and month 11. Note that these criteria match exactly one row. It is also possible to enter a single effort value for multiple rows.

The server replies with a number indicating how many rows were updated in the landing table:

```

<soapenv:Envelope
  xmlns:soapenv="http://schemas.xmlsoap.org/soap/envelope/">
  <soapenv:Body>
    <updateEffortResponse xmlns:ns="http://ws.ceder.jrc.it/xsd">
      <ns:return>1</ns:return>
    </updateEffortResponse>
  </soapenv:Body>
</soapenv:Envelope>

```

2.2.3.3 TAC Prediction Service

The TAC and Quota Prediction service is located at <http://D01RI0700130.net1.cec.eu.int:8080/axis2/services/TacCmdLineService> and is described by a WSDL, which can be obtained by adding the string “?wsdl” at the end of the above URL.

Create forecast

A SOAP client can transmit the following SOAP request, asking the server to create a forecast:

```

<soapenv:Envelope
  xmlns:soapenv="http://schemas.xmlsoap.org/soap/envelope/"
  xmlns:q0="http://ws.ceder.jrc.it/xsd"
  xmlns:q1="http://ws.tac.fishreg.jrc.it/xsd"

```

```

xmlns:q2="http://util.java/xsd"
xmlns:xsd="http://www.w3.org/2001/XMLSchema"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance">
  <soapenv:Body>
    <q0:createForecastArray>
      <q0:bean>
        <q1:area>07.</q1:area>
        <q1:confidencelevel>0.75</q1:confidencelevel>
        <q1:country>GBR</q1:country>
        <q1:lookahead>2</q1:lookahead>
        <q1:minmonthreq>2</q1:minmonthreq>
        <q1:numwinners>2</q1:numwinners>
        <q1:performGraphs>false</q1:performGraphs>
        <q1:species>ANF</q1:species>
        <q1:startfromyear>2001</q1:startfromyear>
      </q0:bean>
    </q0:createForecastArray>
  </soapenv:Body>
</soapenv:Envelope>

```

The above example is asking for a TAC and Quota Prediction for country GBR, species ANF, and area “07.”.

The request above provides the following optional arguments:

Name	Value	Description
Start From	2001	Indicates that data should only be considered at and after 2001. Any data before 2001 is ignored.
Finish at year and month	(none)	Considers all data found in the database, starting from 2001, up to as recent as possible.
Minimal month required	2	Disregard data up to and including February, corresponding to '2'. For all years, only data from March onwards is used.
Predict number of months (“lookahead”)	2	The program will tune the models for 2 months of lookahead, and then will predict the next 2 months from the current data.
Confidence level	0.75	Indicates the prediction confidence level to be used when predicting values. This only works on some models, for

		instance linear models. By default, it will be set to 0.75, which roughly corresponds to “likely”.
Number of winning models	2	Only 2 winning models will perform predictions, which will be returned in a SOAP response.

There is a last parameter, “performGraphs”, which has been set to “false” in the above example. Because SOAP is used for machine-to-machine communication, the TAC and Quota Prediction web service will not generate any graphs.

The server replies with the following SOAP response:

```
<soapenv:Envelope
  xmlns:soapenv="http://schemas.xmlsoap.org/soap/envelope/">
  <soapenv:Body>
    <createForecastArrayResponse
      xmlns:ax25="http://ws.tac.fishreg.jrc.it/xsd"
      xmlns:ax26="http://util.java/xsd"
      xmlns:ns="http://ws.ceder.jrc.it/xsd">
      <ns:return type="it.jrc.fishreg.tac.ws.ForecastAdapter">
        <ax25:area>07.</ax25:area>
        <ax25:batch>20080122.132153.400</ax25:batch>
        <ax25:confidence
          xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
          xsi:nil="true" />
        <ax25:country>GBR</ax25:country>
        <ax25:file
          xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
          xsi:nil="true" />
        <ax25:fitness>0.0181395353188861</ax25:fitness>
        <ax25:forecastId>148</ax25:forecastId>
        <ax25:forecastLines
          type="it.jrc.fishreg.tac.ws.ForecastLineAdapter">
            <ax25:forecastLineId>16103</ax25:forecastLineId>
            <ax25:month>11</ax25:month>
            <ax25:se>0.0272993708253343</ax25:se>
            <ax25:val>0.072193353158329</ax25:val>
```

```

        <ax25:valMax
            xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
            xsi:nil="true" />
        <ax25:valMin
            xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
            xsi:nil="true" />
        <ax25:year>2007</ax25:year>
    </ax25:forecastLines>
    <ax25:forecastLines
        type="it.jrc.fishreg.tac.ws.ForecastLineAdapter">
        <ax25:forecastLineId>16104</ax25:forecastLineId>
        <ax25:month>12</ax25:month>
        <ax25:se>0.0274459822393993</ax25:se>
        <ax25:val>0.0678276820516507</ax25:val>
        <ax25:valMax
            xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
            xsi:nil="true" />
        <ax25:valMin
            xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
            xsi:nil="true" />
        <ax25:year>2007</ax25:year>
    </ax25:forecastLines>
    <ax25:formula>
        arima(dat$var_uptake,order=c(0,0,1),
            seasonal=list(order=c(0,0,1), period=12))
    </ax25:formula>
    <ax25:intervaltype>S</ax25:intervaltype>
    <ax25:knowledgeFromMonth>1</ax25:knowledgeFromMonth>
    <ax25:knowledgeFromYear>2001</ax25:knowledgeFromYear>
    <ax25:knowledgeToMonth>10</ax25:knowledgeToMonth>
    <ax25:knowledgeToYear>2007</ax25:knowledgeToYear>
    <ax25:modelFamily>arima</ax25:modelFamily>
    <ax25:predicting>var_uptake</ax25:predicting>
    <ax25:species>ANF</ax25:species>
</ns:return>
<ns:return type="it.jrc.fishreg.tac.ws.ForecastAdapter">
    <ax25:area>07.</ax25:area>
    <ax25:batch>20080122.132153.400</ax25:batch>

```

```
<ax25:confidence>0.75</ax25:confidence>
<ax25:country>GBR</ax25:country>
<ax25:file
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xsi:nil="true" />
<ax25:fitness>0.0187426645400267</ax25:fitness>
<ax25:forecastId>149</ax25:forecastId>
<ax25:forecastLines
  type="it.jrc.fishreg.tac.ws.ForecastLineAdapter">
  <ax25:forecastLineId>16187</ax25:forecastLineId>
  <ax25:month>11</ax25:month>
  <ax25:se>0.0107481839096373</ax25:se>
  <ax25:val>0.0699442855501644</ax25:val>
  <ax25:valMax>0.091380862423687</ax25:valMax>
  <ax25:valMin>0.0485077086766419</ax25:valMin>
  <ax25:year>2007</ax25:year>
</ax25:forecastLines>
<ax25:forecastLines
  type="it.jrc.fishreg.tac.ws.ForecastLineAdapter">
  <ax25:forecastLineId>16188</ax25:forecastLineId>
  <ax25:month>12</ax25:month>
  <ax25:se>0.0107481839096373</ax25:se>
  <ax25:val>0.102147366060891</ax25:val>
  <ax25:valMax>0.123583942934414</ax25:valMax>
  <ax25:valMin>0.0807107891873689</ax25:valMin>
  <ax25:year>2007</ax25:year>
</ax25:forecastLines>
<ax25:formula>
  lm(var_uptake ~ as.factor(month))
</ax25:formula>
<ax25:intervaltype>B</ax25:intervaltype>
<ax25:knowledgeFromMonth>1</ax25:knowledgeFromMonth>
<ax25:knowledgeFromYear>2001</ax25:knowledgeFromYear>
<ax25:knowledgeToMonth>10</ax25:knowledgeToMonth>
<ax25:knowledgeToYear>2007</ax25:knowledgeToYear>
<ax25:modelFamily>linear</ax25:modelFamily>
<ax25:predicting>var_uptake</ax25:predicting>
<ax25:species>ANF</ax25:species>
```

```

        </ns:return>
      </createForecastArrayResponse>
    </soapenv:Body>
  </soapenv:Envelope>

```

This means that

- An ARIMA modelling variable monthly uptake (not cumulative uptake) has won the modelling contest. It is an ARIMA model of order=c(0,0,1), with seasonal adjustments of order=c(0,0,1), which is a moving average MA(1) model on top of a seasonal moving average SMA(1) model. For months 11 and 12, it predicts quota uptakes of 7.2% and 6.78% respectively.
- Second place went to a linear model where variable monthly uptake is modelled as depending on the month only. For months 11 and 12, it predicts quota uptakes of 6.99% and 10.21% respectively. Based on previous data, it is 75% certain that the November uptake is between 4.85% and 9.13%, and 75% certain that the December uptake is between 8.07% and 12.35%.

Get existing forecast

Using the batch number of “20080122.132153.400”, one can retrieve the same results without having to re-run the prediction process.

```

<soapenv:Envelope
  xmlns:soapenv="http://schemas.xmlsoap.org/soap/envelope/"
  xmlns:q0="http://ws.ceder.jrc.it/xsd"
  xmlns:xsd="http://www.w3.org/2001/XMLSchema"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance">
  <soapenv:Body>
    <q0:getForecastArray>
      <q0:batch>20080122.132153.400</q0:batch>
    </q0:getForecastArray>
  </soapenv:Body>
</soapenv:Envelope>

```

The response is nearly the same as the one for the initial “create forecast”, differences have been marked in **bold underlined**:

```

<soapenv:Envelope
  xmlns:soapenv="http://schemas.xmlsoap.org/soap/envelope/">
  <soapenv:Body>
    <getForecastArrayResponse>

```

```
xmlns:ax25="http://ws.tac.fishreg.jrc.it/xsd"  
xmlns:ax26="http://util.java/xsd"  
xmlns:ns="http://ws.ceder.jrc.it/xsd">  
  <ns:return type="it.jrc.fishreg.tac.ws.ForecastAdapter">
```

... etc ...

```
    </getForecastArrayResponse>  
  </soapenv:Body>  
</soapenv:Envelope>
```


2.2.3.4 Refresh Services

The two services that refresh the prototype's database are the TacFidesCrontLoadService and the TacRefreshInputTablesService.

Please run the two tasks in sequence: First execute the "FIDES CRONT load tables into R data files service", and then execute the "Refresh Input tables from R data files service".

But why do we first fetch the data into some intermediary format, and then refresh the prototype's DB from the intermediary format? This is done because it enables the system to be used "off site", meaning on a computer not located at the European Commission.

In case you plan to run an instance of the prototype system "off site", you will in addition a distribution mechanism for the R data files. We have such a mechanism in place, so please contact us on receiving the proper R data files.

FIDES CRONT load tables into R data files service

The FIDES CRONT load service is located at <http://D01RI0700130.net1.cec.eu.int:8080/axis2/services/TacFidesCrontLoadService> and is described by a WSDL, which can be obtained by adding the string "?wsdl" at the end of the above URL.

A SOAP client can transmit the following SOAP request:

```
<soapenv:Envelope
  xmlns:soapenv="http://schemas.xmlsoap.org/soap/envelope/"
  xmlns:q0="http://ws.ceder.jrc.it/xsd"
  xmlns:xsd="http://www.w3.org/2001/XMLSchema"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance">
  <soapenv:Body>
    <q0:fidesCrontLoad />
  </soapenv:Body>
</soapenv:Envelope>
```

(Note: There is a dummy parameter to fidesCrontLoad that serves no purpose, and can be left empty.)

The server replies with the following SOAP response:

```
<soapenv:Envelope
  xmlns:soapenv="http://schemas.xmlsoap.org/soap/envelope/">
  <soapenv:Body>
    <fidesCrontLoadResponse xmlns:ns="http://ws.ceder.jrc.it/xsd">
      <ns:return/>OK</ns:return>
    </fidesCrontLoadResponse>
  </soapenv:Body>
</soapenv:Envelope>
```

This would mean that the required contents have just been dumped into the R data files.

In case of error, the server will give some information as to why an error occurred:

```
<soapenv:Envelope
  xmlns:soapenv="http://schemas.xmlsoap.org/soap/envelope/">
  <soapenv:Body>
    <fidesCrontLoadResponse xmlns:ns="http://ws.ceder.jrc.it/xsd">
      <ns:return>
        Cannot connect to FIDES CRONT system. Am I on the European Commission's network?
      </ns:return>
    </fidesCrontLoadResponse>
  </soapenv:Body>
</soapenv:Envelope>
```

In case of error, the system does not update the R data files.

Refresh Input tables from R data files service

The Refresh Input tables service is located at <http://D01RI0700130.net1.cec.eu.int:8080/axis2/services/TacRefreshInputTablesService> and is described by a WSDL, which can be obtained by adding the string “?wsdl” at the end of the above URL.

A SOAP client can transmit the following SOAP request:

```
<soapenv:Envelope
  xmlns:soapenv="http://schemas.xmlsoap.org/soap/envelope/"
  xmlns:q0="http://ws.ceder.jrc.it/xsd"
  xmlns:xsd="http://www.w3.org/2001/XMLSchema"
```

```
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance">
  <soapenv:Body>
    <q0:refreshInputTables />
  </soapenv:Body>
</soapenv:Envelope>
```

(Note: There is a dummy parameter to fidesCrontLoad that serves no purpose, and can be left empty.)

The server replies about 5 minutes later, with the following SOAP response:

```
<soapenv:Envelope
  xmlns:soapenv="http://schemas.xmlsoap.org/soap/envelope/">
  <soapenv:Body>
    <refreshInputTablesResponse
      xmlns:ns="http://ws.ceder.jrc.it/xsd">
      <ns:return>OK</ns:return>
    </refreshInputTablesResponse>
  </soapenv:Body>
</soapenv:Envelope>
```

This would mean that the required contents have just been read from the R data files into the prototype's database.

2.3 The database

2.3.1 Introduction

The system is engineered in such a way as to be extensible to different databases. For that, we use a parameter in the program, called the “dbvariant”.

Currently, the sole DB variant recognized is the “postgres-jrc”. This denotes both a dependency on the underlying DBMS (here: postgres) and a dependency on the underlying SQL table structure (here: the one developed at the JRC).

When the code needs to be adapted to a new DB variant, only one module needs to change, so that this new variant is recognized and acted upon.

2.3.2 PostGres JRC Reference Tables

2.3.2.1 Fishery

This table contains the declared landings.

Column	Type	Description
iid	bigserial	Auto-generated ID
regime	varchar(20)	Regime code COCR-QS, COCR-NS
fishery	varchar(10)	Loosely defined fishery code; multiple rows have the same code. “NSF” is Northern Sea Flatfish, “TT” is Tropical Tuna, “SCPEL” is Scottish Pelagic, etc
country	char(3)	ISO country code
species	char(3)	ISO species code
area	varchar(10)	FIDES CRONT area code
comment	varchar(100)	Free text describing the fishery

There can be only one row per area, species, and country.

2.3.3 PostGres JRC Input Tables

2.3.3.1 Landing

This table contains the declared landings.

Column	Type	Description
landing_id	bigserial	Auto-generated ID
area	varchar(16)	FIDES CRONT area code
species	char(3)	ISO species code
year	int4	year of declaration
month	int4	Month at end of which cumulative kilos were declared
country	char(3)	ISO country code
cumul_kgs	float8	Cumulative kilos declared until end of this month for current year
quota_var	float8	quota as it was valid at end of negotiation period
quota_final	float8	quota as it was valid at end year

There can be only one row per area, species, year, month, and country. In other words, the information is aggregated at per – area-species-country level.

2.3.4 PostGres JRC Output Tables

2.3.4.1 Forecast

This table contains the forecast header data.

Column	Type	Description
forecast_id	bigserial	Auto-generated ID
batch	varchar(64)	external process supplied batch number
formula	varchar(128)	R formula such as $\text{lm}(\text{kgs} \sim \text{month} + \text{trend})$
predicting	varchar(16)	What the formula was predicting, meaning the first part of the formula. Possible values are: kgs (kilos declared for current month), cumul_kgs (cumulative kilos declared for current year, up to month), final_uptake (uptake for month with respect to quota_final), cumul_final_uptake (cumulative uptake for current year, with respect to quota_final), var_uptake (uptake for month with respect to quota_var), cumul_var_uptake (cumulative uptake for current year, with respect to quota_var),
model_family	varchar(16)	the kind of formula, which is one of the following: linear, spline, nnet, arima, gam
area	varchar(16)	FIDES CRONT area code
species	char(3)	ISO species code
country	char(3)	ISO country code
knowledge_from_year	int4	year at which model knowledge starts
knowledge_from_month	int4	month at which model knowledge starts (usually 1 for January)
knowledge_to_year	int4	year at which model knowledge stops
knowledge_to_month	int4	month at which model knowledge stops
fitness	float8	Fitness criterion, here the RMSE measure is used. The lower the RMSE, the fitter the model.
intervaltype	char(1)	Value is C, S, B, or N. C means that confidence intervals are given, S means that standard error is given. B means that both C and S are present. N stands for no intervals provided; for instance neural networks do not provide intervals.
confidence	float8	Confidence limit for forecast, usually 95 percent. Only filled if intervaltype is C or B.
file	varchar(2048)	Optional file such as graph of the forecast (deprecated)

There can be only one row per batch, formula, predicting, and model_family.

2.3.4.2 Forecast_line

Column	Type	Description
forecast_line_id	bigserial	Auto-generated ID
forecast_id	int8	Foreign key to forecast table.
year	int4	year to which this particular prediction applies
month	int4	Month to which this particular prediction applies
val	float8	Value being forecast. For the unit of that value, see “forecast.predicting”
val_min	float8	Minimal value forecast, depends on the confidence level. Only filled if “forecast.intervaltype” is C or B.
val_max	float8	Maximal value forecast, depends on the confidence level. Only filled if “forecast.intervaltype” is C or B.
se	float8	Standard error. Only filled if “forecast.intervaltype” is S or B.

There can be only one row per forecast_id, year, and month.

2.3.4.3 Forecast_file

This table is used internally to refer to the files that the user can see on the TAC prediction Results page.

Column	Type	Description
forecast_file_id	bigserial	Auto-generated ID
forecast_id	int8	Foreign key to forecast table.
file_path	varchar(2000)	The directory path name to the graphics file on the local machine.
file_relative_name	varchar(255)	The relative file name of the graphics file on the local machine.
file_type	varchar(31)	A description of the type of file. Current values are “R graph”, depicting how the model fits the data over all years, and “R graph zoom” depicting the fit for the current year.

The file path, the path separator, and the file name together give the absolute file name for a graphics file.

Example: Suppose there were a graphics file named FRA_AE045W_BFT_20071126_142901-(01).png in the directory D:\devel\yms\TacQuotaWeb\r_graphs\

3 Icelandic Prototype System

By: Tryggvi Hjörvar, Fisheries Research Institute, University of Iceland, Reykjavík

3.1 Introduction

3.1.1 Purpose

This document describes the overall design for the Icelandic CEDER Prototype System. The system is primarily aimed at the Atlantic Redfish fisheries, with the objective to demonstrate the feasibility of bringing real-time information on fisheries to stakeholders.

3.1.2 Scope

The name of the system is "CARFI" (Ceder Atlantic Redfish Fisheries Information system).

The system is comprised of three parts:

- Data collection module
- Analysis module
- Presentation module

3.1.2.1 Data collection

The system collects data delivered from the already implemented systems of the Icelandic Oceanographic Institute and the Icelandic Directorate of Fisheries and stores in a database for analysis.

The data collected includes:

- VMS data
 - Date and time
 - Position (longitude, latitude)
 - Target Species
 - Gear type
- Logbook / eLogbook
 - Date of gear deployment
 - Position of gear deployment
 - Catch weight (green weight)
- Landing reports
 - Trip dates and times
 - Landed catch

3.1.2.2 Analysis

The system includes an analysis module whose purpose is twofold.

First, to run classification algorithms on the VMS-data to determine when a vessel is trawling (gear deployed) and when it is cruising. The result is an estimate of effort.

Second, to predict individual vessel and total fleet catch based on effort, by selecting and running an appropriate statistical model from its model library. The prototype system however only includes one prediction model for demonstration purposes.

3.1.2.3 Presentation

Finally, the system includes a presentation module, which is capable of displaying the analysis results in various predefined ways, as static reports or navigable datasets. The prototype system only offers a few predefined reports for demonstration purposes.

3.2 System Overview

The three system modules (data collection, analysis and presentation) each interact with the database and system files as illustrated below.

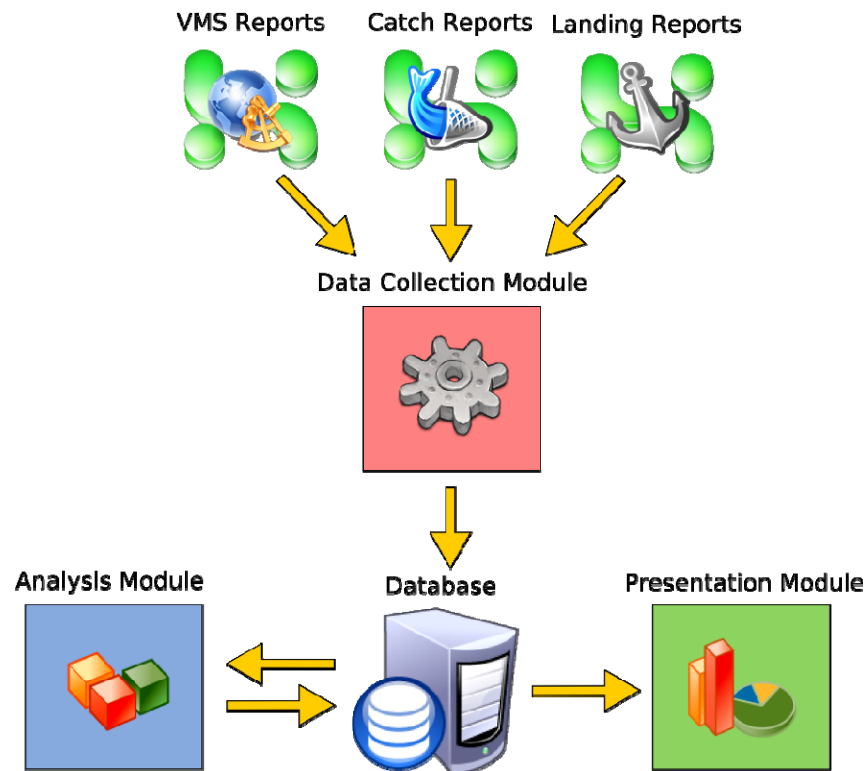


Illustration 1: System overview

The data sources for the data collection module are CSV-formatted files delivered from the Icelandic Fisheries directorate, either to an FTP-site or by e-mail. In either case, the system regularly checks for the presence of new files to process in specific file locations. These files are then imported into the database.

The system includes scripts whose purpose it is to cleanse and conform the data. Running these routines is a prerequisite for any further analysis.

The analysis module is comprised of two independent parts; one for classification of the VMS-track data and effort estimation, and the other for catch prediction based on the effort. The analysis module can incorporate different catch prediction models for other fisheries, but in the prototype system will only work with the Atlantic Redfish model as previously stated.

Finally, the presentation module is written as an active web-page using PHP-scripts on a WAMP (Windows-Apache-MySQL-PHP) platform. It allows the user to examine vessel tracks, catches, estimated effort, predicted catch, and compare in various ways with actual data. It also allows highlighting of interesting vessel behaviour based on the model results.

3.3 Data Collection Module

3.3.1 Import VMS-positions, catch and landing data

Description: The user places a new batch of data in a predefined CSV-format (see below) in a designated file folder. The user can then manually run the import routine via the user interface, or wait for the next scheduled automatic import.

The system import reads the data from the excel-file and inserts into relevant database tables.

Actors: Admin, Directorate of Fisheries or System Scheduler

3.3.2 VMS-positions CSV-file format

The required format for the VMS-position data is a Comma-Separated-Values file. The separator must be a semicolon (;) and the file of the following structure:

Field (column)	Description	Format	Example
Timestamp	Time of position	YYYY.MM.DDWWHH:MM:SS (WW represents whitespace)	2001.02.15 00:12:53
Vessel id	Unique vessel id	Integer number or character code	15
Latitude	Position in degrees	DMS or decimal degrees	N7120
Longitude	Position in degrees	DMS or decimal degrees	-23.53

Each row in the file represents the unique vessel position at the given time.

3.3.3 Catch and landings CVS-file format

The required format for the catch and landing data is a Comma-Separated-Values file. The separator must be a semicolon (;) and the file of the following structure:

Field (column)	Description	Format	Example
Date	Date of catch	DD.MM.YYYY	15.02.2001
Linetime	Length of haul in minutes	Integer	240
Estimated	Catch in kilogrammes as weighed or estimated on deck	Integer	1200
Vessel id	Unique vessel id	Integer number or character code	15
Latitude	Position in degrees where haul started	Decimal degrees without decimal point	6261667
Longitude	Position in degrees where haul started	Decimal degrees without decimal point	-2740000
ICES	ICES-code	Character string	54C2
EEZ	Economic Zone	Character string representing the economic catch zone	island.200.miles
ICESArea	ICES-area	Character string	XIVb
Landed	Officially weighed amount landed in kilogrammes	Integer	472596

Each row in the file represents a catch logbook entry, detailing the position of the start of the haul, the date of the haul, length in minutes, estimated catch, and if this is the last catch of a trip, the officially weighed total when landed.

3.4 Analysis Module

3.4.1 Track Classification and Effort estimation

Description: When the data has been imported the track classification can be run.

The track classification runs scripts to analyse each vessel track and flag them according to likely vessel action. From this the scripts calculate the vessel effort.

Actors: Admin or System Scheduler

3.4.2 Catch Prediction

Description: When the analysis module has finished the effort estimation, the catch prediction scripts can be run. The system calculates the catch prediction for individual vessels and the fleet as a whole and stores the results.

Actors: Admin or System Scheduler

3.5 Presentation Module

3.5.1 View vessel tracks

Description: The individual vessel tracks available in the system can be viewed via a web-based interface. The activity classification of the individual track legs can be colour-coded.

Actors: Authorised users

3.5.2 View fleet activity

Description: The fleet activity, based on vessel tracks and reported gear deployment can be viewed via a web-based interface.

Actors: Authorised users

3.5.3 View predicted catch

Description: The predicted catch for the registered vessels, as well as actual landing data can be viewed in a report via a web-based interface. Discrepancies are highlighted for clarity.

Actors: Authorised users

3.6 Database design

3.6.1 Entity-relationship diagram

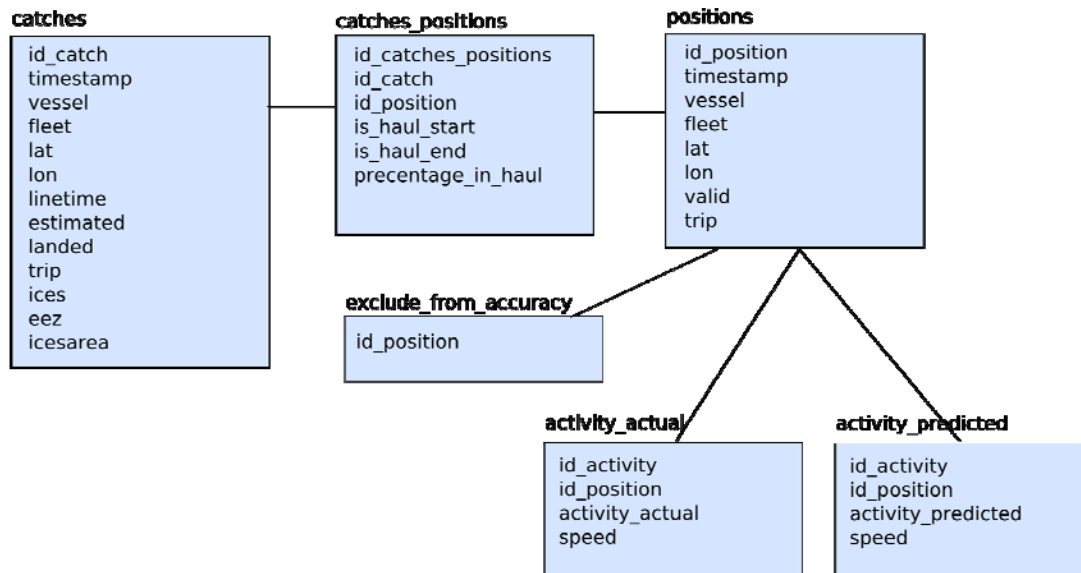


Illustration 2: Database ER-diagram

The catches and positions are connected through a link table, populated by one of the analysis scripts. This is to facilitate analysis of classification errors.

The exclude_from_accuracy table is manually populated with positions that have been deemed unreliable for the purposes of estimating algorithm accuracy.

The analysis module also populates the activity tables, enabling a comparison between classified activity and actual activity.

4 CEDER System for Greenland fisheries

By: Sirius IT

4.1 Introduction

One of the major problems of fish stock assessment is that the process includes significant time-lags as catch data from one year often only is available the year after.

Another problem which causes inaccuracies in the quota-uptake calculation is that the calculations in most countries are based on logbooks and landing declarations which often are inaccurate in respect of the amount caught. We will show that sales-notes can be used to calculate the quota-uptake.

EU has decided (Council Regulation (EC) No 1966/2006) to implement an electronic logbook on all EU fishing vessels above 15 meters within the next 4 years, which will have the consequence that logbook information will be available no later than 24 hours after the fish has been caught. Sirius IT will therefore concentrate the trial on providing more accurate information on catches and discards, thus leading to more accurate and timely quota-uptakes.

The trial will show that we are able to calculate reliable catch data on a daily basis, based on information from hail-messages, logbooks and sales-notes.

4.2 Database Design

4.2.1 General

The trial is based on data from The Greenlandic fishing authorities GFLK, due to this fact the trial will only focus on quota-uptake not on TAC-uptake.

4.3 System design

It is the experience of Sirius IT, that fishing authorities in different countries always have, and will have, different requirements to the database design; we have therefore chosen to design a system that extracts data from the existing database into a few tables. The system is designed in a way that a minimum of effort is required to adapt the system to different database designs.

This path has been chosen in order to be able to use the system in as many different environments as possible.

4.3.1 Input sources of the GFLK database

The database gets input from the following sources:

- Vessel register
- Logbooks including discards
- Sales-notes
- Hail messages
- Vtrack (VMS)
- Observer reports
- Permits
- National quota register

4.3.2 Design of database for catch-control and Quota-management

The information in the above mentioned registers often lies in more than one database at the national fishery administrations, these data has to be harmonized in order to combine information from hail-messages/logbook-register and sales-notes. The task of harmonizing and transforming the data to the data model below, is a task that is not to be underestimated, as each country has different national legislation and different ways of implementing a data-model for the above mentioned registers.

Due to the fact that GFLK is using an Oracle database for the above mentioned registers, the trial software will also be based on an Oracle database, the user interface will be based on the open source product Apache Tomcat, should an administration wish to use another type of database such as PostGres or MySql that will be possible

Information from the hail-message/logbook-register and sales-notes register is transformed to the following tables:

TRIP_SALES

Column Name	Data Type	Description	Source
ID	NUMBER (12)	Primary key	
VESSEL_OWNER_ID ENT	VARCHAR2 (50 Byte)	Name of the vessel owner	Vessel register
VESSEL_IDENT	VARCHAR2 (50 Byte)	Radio callsign of vessel	Vessel register
AUCTION_IDENT	VARCHAR2 (200 Byte)	Name of Auction	Salesnote register
TRIP_NO	VARCHAR2 (8 Byte)	Trip number	Hailmessages/logbook register or salesnote register
SOLD_DATE	DATE	Date of sale	Salesnote register
SELLER_IDENT	VARCHAR2 (200 Byte)	Name of seller	Salesnote register
BUYER_IDENT	VARCHAR2 (200 Byte)	Name of buyer	Salesnote register
LOGBOOK_TYPE	VARCHAR2 (16 Byte)	Type of logbook	Logbook register

CATCH-EFFORTS

Column Name	Data Type	Description	Source
ID	NUMBER (12)	Primary key	
LOG_CATCH_ID	NUMBER (12)	Foreign key to Logbook register	Logbook register
LOGBOOK_IDENT	VARCHAR2 (200 Byte)	Foreign key to Logbook register	Logbook register
GEAR_TEXT	VARCHAR2 (2000 Byte)	Type of gear used	Logbook register
CATCH_DATE	DATE	Day of catch	Logbook register
HAUL_NO	NUMBER (3)	Haul number	Logbook register
HAUL_BEGIN_TIME	DATE	Time when the haul starts	Logbook register
HAUL_END_TIME	DATE	Time when the haul ends	Logbook register
HAUL_VECTOR	SDO_GEOMETRY	Geographical information of the catch	Logbook register
INT_AREA_NAME	VARCHAR2 (20 Byte)	Name of the Quota area	Logbook register
INT_AREA_SQUARE	VARCHAR2 (20 Byte)	name of the sub area	Logbook register
PERMIT_IDENT	VARCHAR2 (50 Byte)	Name of permit	Permit register
PERMIT_FLAG_STATE	VARCHAR2 (3 Byte)	Flagstate of permit	Permit register
PERMIT_AREA_NAME	VARCHAR2 (20 Byte)	Name of area where the permit is valid	Permit register
SUB_SPECIES_CODE	VARCHAR2 (50 Byte)	Sub specie code	Logbook register
CATCH_SIZE_CAT	VARCHAR2 (50 Byte)	Size category the specie	Logbook register
SPECIES_CODE	VARCHAR2 (20 Byte)	Code for the specie	Logbook register
AMOUNT_KG	NUMBER (12,3)	Amount to be deducted from the quota	Calculated quota uptake
VALUE_LOCAL_CURR	NUMBER (12,3)	Value of the sold specie	Salesnote register

TRIP_SALES_ID	NUMBER (12)	Foreign key	
LOG_AMOUNT_KG	NUMBER (12,3)	Amount as reported in the logbook	Hailmessages/Logbook register
SALES_AMOUNT_KG	NUMBER (12,3)	Amount as reported in the salesnote	Salesnote register

4.3.3 Catch and quota calculation

At any given time it is possible to calculate the quota-uptake for relevant species.

The flow of information in the GFLK system is as follows:

The vessel transmits hail-messages on a weekly basis (vessels on shorter trips does not transmit hail-messages)

Logbooks are sent to the authorities when the fishing trip has ended.

Sales-notes are sent to the authorities when the fish has been sold.

As soon as the hail message is entered in the system the reported catch will be deducted from the quota, when the fishing trip has ended and the logbook data has been entered in the database the logbook data will replace the catch data from the previous received hail messages. When the fish has been sold and the sales-note data has been entered in the database the sales-note data replaces the logbook data except for the discards which is retained from the logbook.

Note:

When the electronic logbook has been implemented, it will be possible to use logbook data on a daily basis, as the fisherman will be obliged to transmit logbook data at least every 24 hours.

4.3.4 Data-mining

Apart from quota uptake the system can be used for data-mining in order to look for anomalies. E.g. Alarms will be raised if for example catches from a given area are above the 90 percentile.

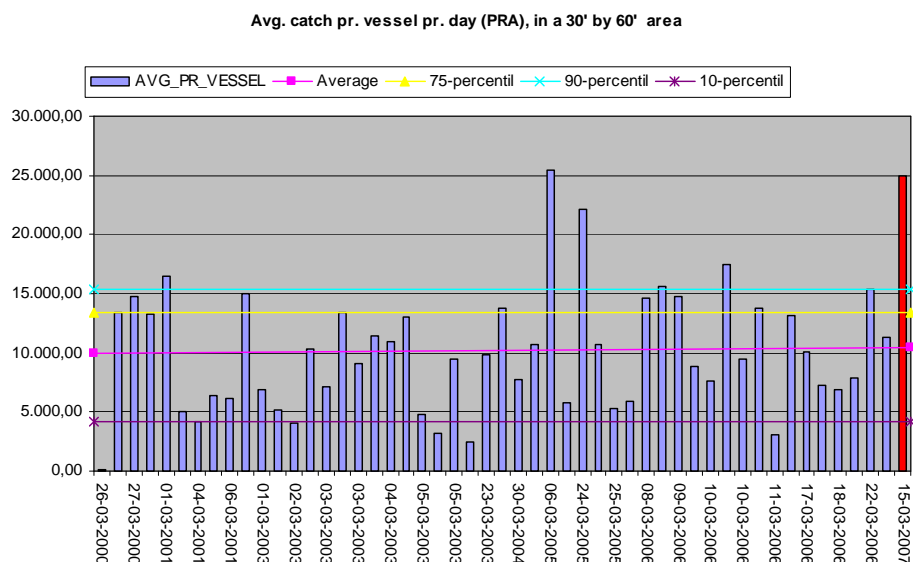


Figure 1 on the left the difference between logbooks and sales notes is plotted. On the right the catches per vessel per day in a particular area are shown

(Page intentionally left blank)

5 Correlation Systems: CEDER Web Site

By: Noa Dayagi, Correlation Systems

Data contributed from: A. Kideys (JRC), J. Neves (NEAFC), F. Quirijns (IMARES), G. Pilling (Cefas), David Reid (FRS).

5.1 Top Level Design

5.1.1 General

The CEDER system is implemented as a web site. The web site enables the user to perform queries, view reports and manually insert data (such as VMS data, electronic logbooks, etc.) into the system's database. A future extension to the system will include the proposed analyses as web services and as stand alone web services.

5.1.2 The System Design

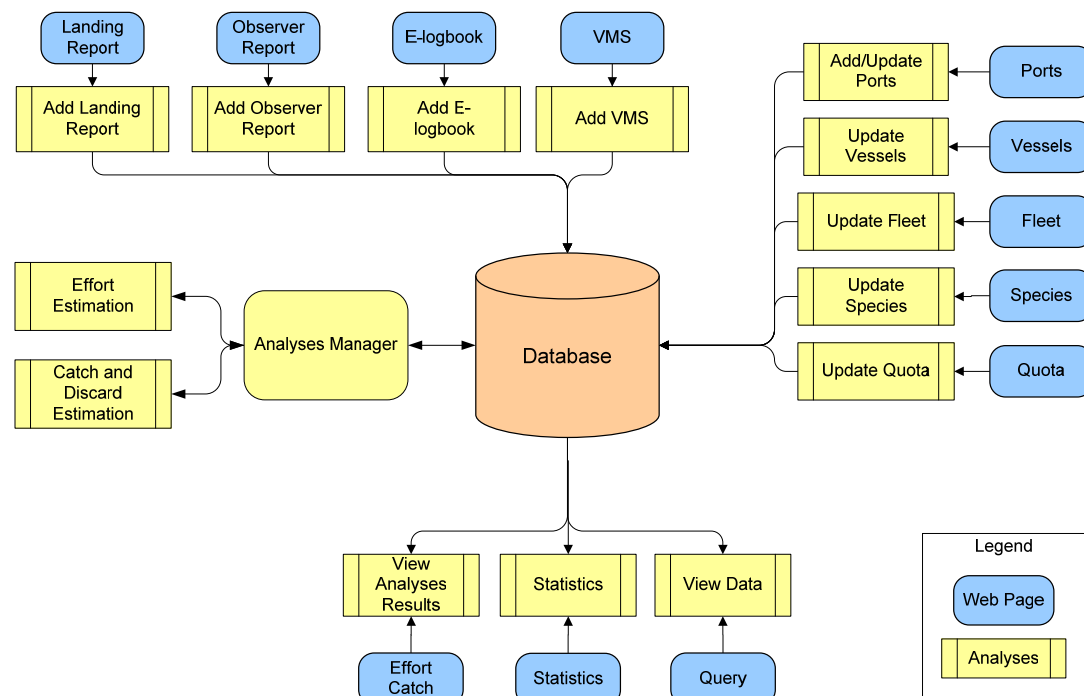


Figure 2 – Top Level Design

5.1.2.1 Web Site

The Web Site, through its different pages, supports all of the available analyses.

The Web Site supports 4 input data types:

1. VMS data
2. E-logbooks
3. Observer Reports
4. Landing Reports

The data can be inserted into the system's database using the Add analyses. The user can also update ports details, vessel and fleet details and species details which are also maintained in the database. The user can view the data stored in the database using the View Data pages.

In addition the web site enables the user to perform analyses on the stored data, such as:

1. **Effort Estimation** – the analysis is performed on VMS data of a single vessel, and determines for each record whether the vessel was fishing, cruising or at the port. The records are collected into tracks, which are displayed on a map along with the Effort estimation for each point.
2. **Catch and Discard Estimation** – the analysis is performed on all data types which are stored in the database. The analysis estimates for a specified area/ species and country the weight of catch and discard, and the catch/quota ratio. The output is displayed as a data table and on charts on the web pages.

The user can also perform statistics on the results of the Effort Estimation analysis.

5.1.2.2 Web Services

A future extension to the analyses above will enable the user to insert data into the database and retrieve data from the database using Web Services. The user will also be able to perform the analyses on the data using Web Services.

5.1.2.3 Stand Alone Web Services

Another future extension will include the Stand-Alone web services, which are independent services that don't use the system's database, and which enable external users to utilize products of the project, for their research and other needs. These services will reveal to external users the algorithms which were developed during the project, without forcing them to insert their data to the system's database. That way, the users will be able to use their own database. The stand-alone web services will include the effort estimation algorithm, and the catch and discard algorithm.

5.2 The Web Site – Detailed Design

5.2.1 Input Analyses

The input analyses are used to upload data into the system's database, such as: VMS data, e-logbooks, observer reports and landing reports. The uploading of the data into the database is performed manually, using the web site to select a file, and uploads it into the database.

5.2.1.1 Add VMS

5.2.1.1.1 Description

The Add VMS analysis enables the user to add VMS records to the system's database.

5.2.1.1.2 Input

The input consists of a file of VMS records, each record is in the VMS message format:

Field Name	Data Type	Description	Units	Accuracy	Optional
Vessel identifier	string	The call sign of the vessel	N/A	N/A	No
Country	string	The country's name (3 letters)	N/A	N/A	No
Fleet Name	string	The name of the fleet	N/A	N/A	No
Detection time	DateTime	The date and time of the record	Month/day/year hour:min:sec AM/PM	N/A	No
Latitude	double	The Y coordinate of the vessel's location	WGS-84	1 meter	No
Longitude	double	The X coordinate of the vessel's location	WGS-84	1 meter	No
Speed	double	Instantaneous speed of the vessel	Kmh -1 if not reported	1 kmh	Yes
Heading	double	The estimated direction of the vessel's progress	Degrees from the north -1 if not reported	0.1 degree	Yes
Reported Activity	Enum	Effort which was reported by the vessel	0 = Unknown 1 = Fishing 2 = cruising 3 = harbor 4 = other	N/A	Yes

5.2.1.1.3 Output

The output is displayed as a message on the screen, indicating whether the process was successfully completed, or has failed.

5.2.1.2 Add E-Logbook

5.2.1.2.1 Description

The Add E-Logbook analysis enables the user to add an E-logbook to the system's database.

5.2.1.2.2 Input

The input is an E-logbook file, each record is in the E-logbook format:

Field Name	Data Type	Description	Units	Accuracy	Optional
Vessel identifier	string	The call sign of the vessel	N/A	N/A	No
Country	string	The country's name (3 letters)	N/A	N/A	No
Fleet Name	string	The name of the fleet	N/A	N/A	No
Start Time	DateTime	The start time of the track	Month/day/year hour:min:sec AM/PM	N/A	No
End Time	DateTime	The end time of the track	Month/day/year hour:min:sec AM/PM	N/A	No
Area	string	The area in which the location was recorded	N/A	N/A	No
Catch species abbreviation	string	The abbreviation of the species that was caught	N/A	N/A	No
Catch Weight	double	The weight of the species that was caught	Tons	1 kg	No
Discard species abbreviation	string	The abbreviation of the species that was discarded	N/A	N/A	Yes
Discard Weight	double	The weight of the species that was discarded	Tons	1 kg	No
Reported Activity	Enum	Effort which was reported by the vessel	0 = Unknown 1 = Fishing 2 = cruising 3 = harbor 4 = other	N/A	Yes

5.2.1.2.3 Output

The output is displayed as a message on the screen, indicating whether the process was successfully completed, or has failed.

5.2.1.3 Add Observer Report

5.2.1.3.1 Description

The Add Observer Report analysis enables the user to add an Observer Report file to the system's database.

5.2.1.3.2 Input

The input is an observer -report file, each record is in the Observer Report format:

Field Name	Data Type	Description	Units	Accuracy	Optional
Vessel identifier	string	The call sign of the vessel	N/A	N/A	No
Country	string	The country's name (3 letters)	N/A	N/A	No
Fleet Name	string	The name of the fleet	N/A	N/A	No
Start Time	DateTime	The start time of the track	Month/day/year hour:min:sec AM/PM	N/A	No
End Time	DateTime	The end time of the track	Month/day/year hour:min:sec AM/PM	N/A	No
Catch species abbreviation	string	The abbreviation of the species that was caught	N/A	N/A	No
Catch Weight	double	The weight of the species that was caught	Tons	1 kg	No

Discard species abbreviation	string	The abbreviation of the species that was discarded	N/A	N/A	Yes
Discard Weight	double	The weight of the species that was discarded	Tons	1 kg	No
Reported Activity	Enum	Effort which was reported by the vessel	0 = Unknown 1 = Fishing 2 = cruising 3 = harbor 4 = other	N/A	Yes

5.2.1.3.3 Output

The output is displayed as a message on the screen, indicating whether the process was successfully completed, or has failed.

5.2.1.4 Add Landing Report

5.2.1.4.1 Description

The Add Landing Report analysis enables the user to add a Landing Report file to the system's database.

5.2.1.4.2 Input

The input is a Landing Report file, each record is in the Landings format of DG-FISH. Each record of the monthly landing report consists of the following fields:

Field Name	Data Type	Description	Units	Accuracy	Optional
Year	int	The year if the record	N/A	N/A	No
Month	int	The month of the record	N/A	N/A	No
Area	string	The area in which the location was recorded	N/A	N/A	No
Country	string	The country's name (3 letters)	N/A	N/A	No
Species abbreviation	string	The abbreviation of the species in the area	N/A	N/A	No
Catch weight	double	The net catch weight of the specific species	Tons	1 kg	No

5.2.1.4.3 Output

The output is displayed as a message on the screen, indicating whether the process was successfully completed, or has failed.

5.2.2 Parameters Update

The Parameters Update pages are used to insert and update details of the different parameters which are used by the various analyses of the system.

5.2.2.1 Add Vessel characteristics

5.2.2.1.1 Description

The Add Vessel Characteristics analysis enables the user to insert a vessel characteristic to the system's database.

5.2.2.1.2 Input

The input consists of a list of vessel's parameters; each record of a vessel consists of the following fields:

Field Name	Data Type	Description	Units	Accuracy	Optional
Vessel Identifier	string	The call sign of the vessel	N/A	N/A	No
Country	string	The country's name (3 letters)	N/A	N/A	No
Engine power	double	The vessel's engine power	Kw	N/A	No
Length	double	The vessel's length	Meters	1 meter	No
Fleet Name	string	The name of the fleet	N/A	N/A	No

5.2.2.1.3 Output

The output is displayed as a message on the screen, indicating whether the process was successfully completed, or has failed.

5.2.2.2 Add Fleet characteristics

5.2.2.2.1 Description

The Add Fleet Characteristics analysis enables the user to insert a fleet characteristic to the system's database.

5.2.2.2.2 Input

The input consists of a list of fleet names:

Field Name	Data Type	Description	Units	Accuracy	Optional
Fleet Name	string	The name of the fleet	N/A	N/A	No
Country	string	The country's name (3 letters)	N/A	N/A	No
Gear type	string	The vessel's gear	N/A Default: Unknown	N/A	No
Mesh size	double	The vessel's mesh size	N/A	N/A	No
Species abbreviation	string	The name of the vessel's target specie	N/A	N/A	No
Species Percentage	string	The percentage of the species from the total catch weight	Between 0 and 100	N/A	No
Max Trip Time	int	Maximum days at sea	Days Default: -1	N/A	Yes
Min Speed	double	The minimum speed of the vessel	Kmh Default: -1	N/A	Yes
Max Speed	double	The maximum speed of the vessel	kmh Default: -1	N/A	Yes
Min Speed Std	double	The minimum speed std of the vessel	Default: -1	N/A	Yes
Max Speed Std	double	The maximum speed std of the vessel	Default: -1	N/A	Yes
Min Density	double	The minimum density of the vessel	Km Default: -1	N/A	Yes
Max Density	double	The maximum density of the vessel	Km Default: -1	N/A	Yes
Density Threshold	double	The threshold of the density parameter of the vessel	Km Default: -1	N/A	Yes
Min Speed Individual Threshold For High Sampling rate	double	The minimum speed threshold for high sampling rate of the vessel	KmH Default: -1	N/A	Yes
Max Speed Individual Threshold For High Sampling rate	double	The maximum speed threshold for high sampling rate of the vessel	KmH Default: -1	N/A	Yes
Min Speed Individual	double	The minimum speed threshold for	KmH	N/A	Yes

Threshold For Low Sampling rate		low sampling rate of the vessel	Default: -1		
Max Speed Individual Threshold For Low Sampling rate	double	The maximum speed threshold for low sampling rate of the vessel	KmH Default: -1	N/A	Yes
Speed Mean Threshold	double	The speed mean threshold of the vessel	KmH Default: -1	N/A	Yes
Speed Std Threshold	double	The speed std threshold of the vessel	Default: -1	N/A	Yes
Min Distance	double	The minimum distance of the vessel	Km Default: -1	N/A	Yes
Max Distance	double	The maximum distance of the vessel	Km Default: -1	N/A	Yes
Distance Threshold	double	The distance threshold of the vessel	Km Default: -1	N/A	Yes
Number of vessels in the fleet	int	The number of vessels that behave as the vessels	N/A	N/A	Yes

5.2.2.2.3 Output

The output is displayed as a message on the screen, indicating whether the process was successfully completed, or has failed.

5.2.2.3 Add Ports

5.2.2.3.1 Description

The Add Ports analysis enables the user to insert Port details to the system's database.

5.2.2.3.2 Input

The input consists of the port's details. Each record of a port consists of the following fields:

Field Name	Data Type	Description	Units	Accuracy	Optional
Port Name	string	The name of the port	N/A	N/A	No
Country	string	The country's name (3 letters)	N/A	N/A	No
Latitude	double	The Y coordinate of the port's location	WGS-84	1 meter	No
Longitude	double	The X coordinate of the port's location	WGS-84	1 meter	No

5.2.2.3.3 Output

The output is displayed as a message on the screen, indicating whether the process was successfully completed, or has failed.

5.2.2.4 Add Species

The Add Species analysis can be used only by a user which is an administrator.

5.2.2.4.1 Description

The Add Species analysis enables the user to insert species details to the system's database.

5.2.2.4.2 Input

The input consists of a list of species details, as described in FAO Species list. Each record consists of the following fields:

Field Name	Data Type	Description	Units	Accuracy	Optional
Species Name	string	The name of the species	N/A	N/A	No
Species Abbreviation	string	The abbreviation of the species	N/A	N/A	No

5.2.2.4.3 Output

The output is displayed as a message on the screen, indicating whether the process was successfully completed, or has failed.

5.2.2.5 Add Quota

5.2.2.5.1 Description

The Add Quota analysis enables the user to insert the quota for each species in a specific year to the system's database.

5.2.2.5.2 Input

The input consists of a list of species and their quota, each record consists of the following fields:

Field Name	Data Type	Description	Units	Accuracy	Optional
Country	string	The country's name (3 letters)	N/A	N/A	No
Year	int	The year relevant for the data	N/A	N/A	No
Area	string	The area in which the location was recorded	N/A	N/A	Yes
Species Abbreviation	string	The abbreviation of the species	N/A	N/A	No
Quota	double	The quota defined for the species	Tons	1 Ton	No

5.2.2.5.3 Output

The output is displayed as a message on the screen, indicating whether the process was successfully completed, or has failed.

5.2.3 Processing Analyses

The processing analyses enable the user to analyze the data which is stored in the database. The available analyses are: Effort estimation and Catch and Discard estimation.

5.2.3.1 Effort Estimation

5.2.3.1.1 Description

The Effort Estimation analysis returns the estimated fishing activity for each VMS record of a specified vessel, based on the VMS data. The possible fishing activities are: fishing, cruising, or at port.

The Effort Estimation analysis generates tracks from the VMS records, and then performs the Effort Estimation on each track separately.

5.2.3.1.2 Input

The input of the analysis consists of a vessel name, which exists in the database, along with a fleet name and a country and a time frame. The input for the Effort analysis consists of the following fields:

Field Name	Data Type	Description	Units	Accuracy	Optional
------------	-----------	-------------	-------	----------	----------

Vessel identifier	string	The call sign of the vessel	N/A	N/A	No
Fleet Name	string	The name of the fleet	N/A	N/A	No
Country	string	The country's name (3 letters)	N/A	N/A	No
Start Time	DateTime	The start time of the track	Month/day/year hour:min:sec AM/PM	N/A	No
End Time	DateTime	The end time of the track	Month/day/year hour:min:sec AM/PM	N/A	No

5.2.3.1.3 Output

The output of the analysis includes tracks, which consist of segments, which consist of VMS records. Each record of a segment represents a change in the estimated Effort. Each track can be displayed on the map, each segment in a colour which represents the fishing activity it contains.

Each track consists of the following fields:

Field Name	Data Type	Description	Units	Accuracy	Optional
List of Segments	Segment Collection	List of segments	N/A	N/A	No
List of Parameters	Parameter Collection	Parameters of the track	N/A	N/A	No

5.2.3.1.3.1 A Segment

Each segment consists of the following fields:

Field Name	Data Type	Description	Units	Accuracy	Optional
Start Time	DateTime	the start date and time of the record	Month/day/year hour:min:sec AM/PM	N/A	No
End Time	DateTime	The end date and time of the record	Month/day/year hour:min:sec AM/PM	N/A	No
Start Longitude	double	The X coordinate of the segment's start location	WGS-84	1 meter	No
Start Latitude	double	The Y coordinate of the segment's start location	WGS-84	1 meter	No
End Longitude	double	The X coordinate of the segment's end location	WGS-84	1 meter	No
End Latitude	double	The Y coordinate of the segment's end location	WGS-84	1 meter	No
Speed	double	instantaneous speed of the vessel	Kmh	1 kmh	No
Heading	double	The estimated direction of the vessel's progress	Degrees from the north	0.1 degree	No
Estimated Activity	Enum	Effort which was estimated by the Effort estimation algorithm	0 = Unknown 1 = Fishing 2 = cruising 3 = harbor 4 = other	N/A	No
Parameters	Parameter Collection	Parameters of the segment	N/A	N/A	No

5.2.3.2 Catch and Discard Estimation

5.2.3.2.1 Description

FP6 CEDER Project Deliverable 2.1 Design of systems to be used for pilot.
Software architecture of the CEDER project's pilot systems.

The Catch and Discard Estimation analysis returns the estimated catch and discard for a specified area, species or country. The analysis is performed on all data types which include relevant data.

5.2.3.2.2 Input

The input of the analysis can consist of parameters such as a country, a species name or an area - which exists in the database, and a time frame. The input for the Catch and Discard analysis consists of the following fields:

Field Name	Data Type	Description	Units	Accuracy	Optional
Year	int	The desired year	N/A	N/A	No
Country	string	The country's name (3 letters)	N/A	N/A	Yes
Species Abbreviation	string	The abbreviation of the species in the area	N/A	N/A	Yes
Area	string	The desired area	N/A	N/A	Yes
Quota	double	The amount of quota to filter	Tons	1 Ton	No

5.2.3.2.3 Output

The output of the analysis includes the estimated weight of catch and discard, and the catch/quota ratio. The output is displayed on a data table, on charts and on the map. The output shall consist of the following fields, according the input parameters that were selected:

Field Name	Data Type	Description	Units	Accuracy	Optional
Year	int	The desired year	N/A	N/A	No
Month	int	The months of the selected years	N/A	N/A	No
Catch	double	The catch weight of the selected parameters	Tons	1 kg	Yes
Discard	double	The weight of discard during the catch trips of the selected parameters	Tons	1 kg	Yes
Quota	double	The weight's quota of the selected parameters	Tons	1 kg	Yes
Catch/Quota	double	The ratio catch/quota	Percent	0.01 percent	Yes

5.2.4 Query Analyses

The query analyses are used to view data which is stored in the system's database, such as viewing details on vessels, ports, areas, countries, fleets and species.

5.2.4.1 Fleet Query

The Fleet query enables the user to retrieve the fleets which are stored in the database. This query is performed automatically when the page is loaded, and the results are displayed in a table on the screen.

5.2.4.2 Vessels Query

The Vessels query enables the user to retrieve the vessels which are stored in the database. The parameters for the query are: country, fleet name and target species. The results are displayed in a table on the screen.

5.2.4.3 Ports Query

The Ports query enables the user to retrieve the ports details which are stored in the database. This query is performed automatically when the page is loaded, and the results are displayed in a table on the screen. The ports can be displayed on the map as well.

5.2.4.4 Species Query

The Species query enables the user to retrieve the species details which are stored in the database. This query is performed automatically when the page is loaded, and the results are displayed in a table on the screen.

5.2.4.5 Country Query

The Country query enables the user to retrieve the countries details which are stored in the database. This query is performed automatically when the page is loaded, and the results are displayed in a table on the screen.

5.2.4.6 Areas Query

The Areas query enables the user to retrieve the areas details which are stored in the database. This query is performed automatically when the page is loaded, and the results are displayed in a table on the screen. The areas can be displayed on a map as well.

5.2.5 Statistics

The Statistics screens are used to display the statistics on the output of the processing analyses, such as effort estimation analysis.

5.2.5.1 Effort Statistics

5.2.5.1.1 Description

The Effort Statistics query enables the user to perform statistics on the effort estimation of the vessels, which are stored in the database.

5.2.5.1.2 Input

The input consists of parameters such as: area, country and target species.

Field Name	Data Type	Description	Units	Accuracy	Optional
Start Time	DateTime	The start time of the track	Month/day/year hour:min:sec AM/PM	N/A	No
End Time	DateTime	The end time of the track	Month/day/year hour:min:sec AM/PM	N/A	No
Country	string	The country's name (3 letters)	N/A	N/A	No
Fleet Name	string	The name of the fleet	N/A	N/A	No
Area	string	The area in which the location was recorded	N/A	N/A	No

5.2.5.1.3 Output

The output of the analysis is the number of records that were estimated as fishing, and the number of records that were estimated as cruising, within the specified time frame. This data is displayed on a chart, and the selected area is displayed on the map.

(end of document)

European Commission

EUR 24019 EN – Joint Research Centre – Institute for the Protection and Security of the Citizen

Title: FP6 CEDER Project Deliverable 2.1: "Design of systems to be used for pilot"

Authors: N. Dayagi (Corr.), M. Grønfeldt (Sirius), T. Hjörvar (FRI), A. Kideys (JRC),
U. Kröner (JRC), J. Neves (NEAFC), F. Quirijns (IMARES), G. Pétursdóttir (FRI),
G. Pilling (Cefas), D. Reid (FRS), E. Rosenberg (Corr.), K. Sünksen (GINR)

Luxembourg: Office for Official Publications of the European Communities

2009 – 62 pp. – 21 x 29,7 cm

EUR – Scientific and Technical Research series – ISSN 1018-5593

ISBN 978-92-79-13739-6

DOI 10.2788/40321

Abstract

Addressing the uncertainties in fishing activities, the FP6 CEDER project designed a set of pilot systems to demonstrate the feasibility of providing information in a timelier manner. This document describes the software designs of the aforementioned pilot systems.

How to obtain EU publications

Our priced publications are available from EU Bookshop (<http://bookshop.europa.eu>), where you can place an order with the sales agent of your choice.

The Publications Office has a worldwide network of sales agents. You can obtain their contact details by sending a fax to (352) 29 29-42758.

The mission of the JRC is to provide customer-driven scientific and technical support for the conception, development, implementation and monitoring of EU policies. As a service of the European Commission, the JRC functions as a reference centre of science and technology for the Union. Close to the policy-making process, it serves the common interest of the Member States, while being independent of special interests, whether private or national.



LB-NA-24019-EN-C

ISBN 978-92-79-13739-6

